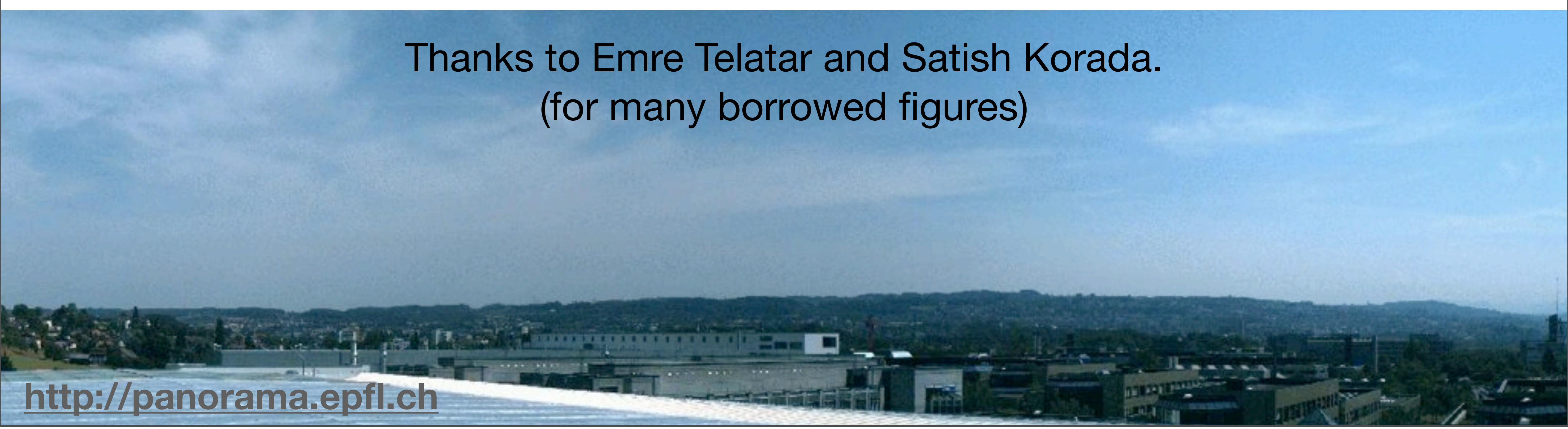
Polar Codes -- A New Paradigm for Coding

R. Urbanke, EPFL

Physics of Algorithms, Santa Fe, September 2nd, 2009



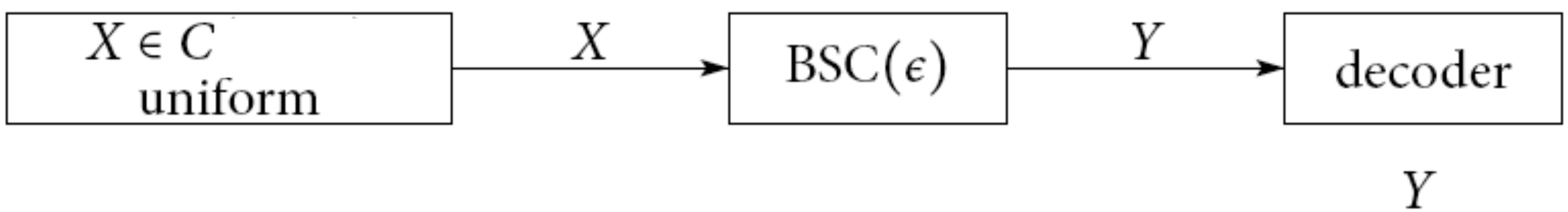
Thanks to Emre Telatar and Satish Korada.
(for many borrowed figures)

http://panorama.epfl.ch

Coding

Coding

Coding



C={000, 010, 101, 111}   code

n ... blocklength

Important Parameters

Important Parameters

$(r, \mathrm{P}, \chi_E, \chi_D, n)$

rate, error probability,
encoding complexity,
decoding complexity,
blocklength

Linear Codes

Linear Codes

$$C(G) = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\}$$

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

generator matrix

Linear Codes

$$C(G) = \{x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k\} \qquad C = \{x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k\} = \{x \in \mathbb{F}^n : Hx^T = 0^T\}$$

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

generator matrix

parity-check matrix

Linear Codes

$$C(G) = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\} \qquad C = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\} = \left\{ x \in \mathbb{F}^n : Hx^T = 0^T \right\}$$

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

generator matrix

parity-check matrix

Bitwise MAP Decoding

$$
\begin{aligned}
\hat{x}_i^{\mathrm{MAP}}(y) &= \mathrm{argmax}_{x_i \in \{\pm 1\}} p_{X_i \mid Y}(x_i \mid y) \\
\text{(law of total probability)} \quad &= \mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} p_{X \mid Y}(x \mid y) \\
\text{(Bayes')} \quad &= \mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} p_{Y \mid X}(y \mid x) p_X(x) \\
\text{(2.13)} \quad &= \mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} \Big(\prod_j p_{Y_j \mid X_j}(y_j \mid x_j) \Big) \mathbb{1}_{\{x \in C\}},
\end{aligned}
$$

Bitwise MAP Decoding

$$\hat{x}_i^{\mathrm{MAP}}(y) = \mathrm{argmax}_{x_i \in \{\pm 1\}} p_{X_i \mid Y}(x_i \mid y)$$

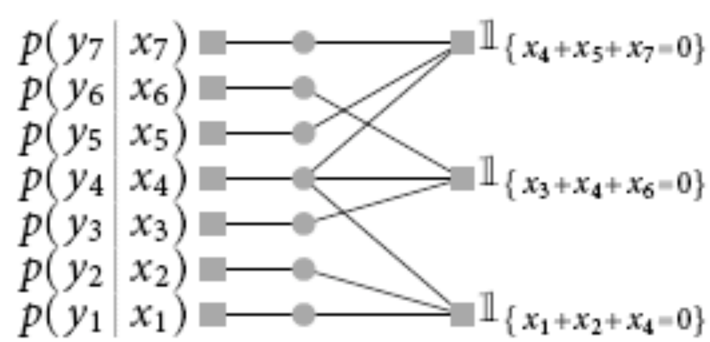
(law of total probability) $$= \mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} p_{X \mid Y}(x \mid y)$$

(Bayes’) $$= \mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} p_{Y \mid X}(y \mid x) p_X(x)$$

(2.13) $$= \mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} \Big(\prod_j p_{Y_j \mid X_j}(y_j \mid x_j) \Big) \mathbb{1}_{\{x \in C\}},$$

$$\mathrm{argmax}_{x_i \in \{\pm 1\}} \sum_{\sim x_i} \Big(\prod_{j=1}^{7} p_{Y_j \mid X_j}(y_j \mid x_j) \Big) \mathbb{1}_{\{x_1+x_2+x_4=0\}} \mathbb{1}_{\{x_3+x_4+x_6=0\}} \mathbb{1}_{\{x_4+x_5+x_7=0\}}$$

$$H = \begin{array}{c} \begin{array}{ccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix} \end{array}$$



[LDPC -- Gallager ‘60]

Polar Codes: Summary

Polar Codes: Summary

Erdal Arikan, ISIT 2007

Polar Codes: Summary

Erdal Arikan, ISIT 2007

very general phenomenon

Polar Codes: Summary

Erdal Arikan, ISIT 2007

very general phenomenon

information theoretic view why codes work

Polar Codes: Summary

Erdal Arikan, ISIT 2007

very general phenomenon

information theoretic view why codes work

first “low complexity” scheme which provably achieves the capacity for a fairly wide array of channels

Polar Codes: Summary

Erdal Arikan, ISIT 2007

very general phenomenon

information theoretic view why codes work

first “low complexity” scheme which provably achieves the capacity for a fairly wide array of channels

many possible variations on the theme

Polar Codes: Summary

Erdal Arikan, ISIT 2007

very general phenomenon

information theoretic view why codes work

first “low complexity” scheme which provably achieves the capacity for a fairly wide array of channels

many possible variations on the theme

codes not only good for channel coding; work equally well for source coding and more complicated scenarios

References

References

arXiv:0807.3917 [ps, pdf, other]

Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels

Erdal Arikan

Comments: 49 pages, 14 figures. Submitted to IEEE Transactions on Information Theory, Oct. 2007

Subjects: **Information Theory (cs.IT)**

arXiv:0807.3806 [ps, pdf, other]

On the Rate of Channel Polarization

Erdal Arikan, Emre Telatar

Comments: 10 pages

Subjects: **Information Theory (cs.IT)**

arXiv:0901.2370 [ps, pdf, other]

Polar Codes for Channel and Source Coding

Nadine Hussami, Satish Babu Korada, Rudiger Urbanke

Comments: submitted to ISIT

Subjects: **Information Theory (cs.IT)**

arXiv:0901.0536 [ps, pdf, other]

Polar Codes: Characterization of Exponent, Bounds, and Constructions

Satish Babu Korada, Eren Sasoglu, Rudiger Urbanke

Comments: Submitted to IEEE Transactions on Information Theory, minor updates

Subjects: **Information Theory (cs.IT)**

arXiv:0901.2207 [ps, pdf, other]

Performance and Construction of Polar Codes on Symmetric Binary-Input Memoryless Channels

Ryuhei Mori, Toshiyuki Tanaka

Comments: 5 pages, 3 figures, submitted to ISIT2009

Subjects: **Information Theory (cs.IT)**

arXiv:0903.0307 [ps, pdf, other]

Polar Codes are Optimal for Lossy Source Coding

Satish Babu Korada, Rudiger Urbanke

Comments: 15 pages, submitted to Transactions on Information Theory

Subjects: **Information Theory (cs.IT)**

arXiv:0907.3291 [ps, pdf, other]

The Compound Capacity of Polar Codes

S. Hamed Hassani, Satish Babu Korada, Ruediger Urbanke

Comments: 5 pages

Subjects: **Information Theory (cs.IT)**

arXiv:0908.0302 [ps, pdf, other]

Polarization for arbitrary discrete memoryless channels

Eren Sasoglu, Emre Telatar, Erdal Arikan

Comments: 12 pages

Subjects: **Information Theory (cs.IT)**

Codes from Kronecker Product of G_2

Codes from Kronecker Product of G_2

$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Codes from Kronecker Product of G_2

$$G_2^{\otimes 2} = \begin{bmatrix} G_2 & 0 \\ G_2 & G_2 \end{bmatrix}$$

Codes from Kronecker Product of G_2

$$G_2^{\otimes 2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Codes from Kronecker Product of G$_2$

$$G_2^{\otimes 3} = \begin{bmatrix} G_2 & 0 & 0 & 0 \\ G_2 & G_2 & 0 & 0 \\ G_2 & 0 & G_2 & 0 \\ G_2 & G_2 & G_2 & G_2 \end{bmatrix}$$

Codes from Kronecker Product of G_2

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Codes from Kronecker Product of G_2

length $N = 2^m$, $m \in \mathbb{N}$

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Codes from Kronecker Product of G_2

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

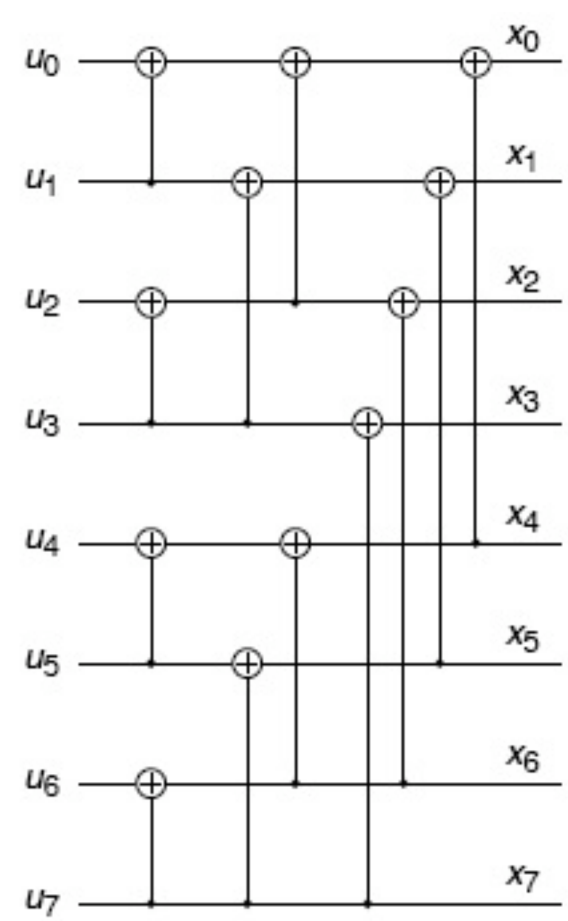
$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Codes from Kronecker Product of G_2

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$



$$\bar{x} = [u_0\ u_1\ u_2\ u_3\ u_4\ u_5\ u_6\ u_7] G_2^{\otimes 3}$$

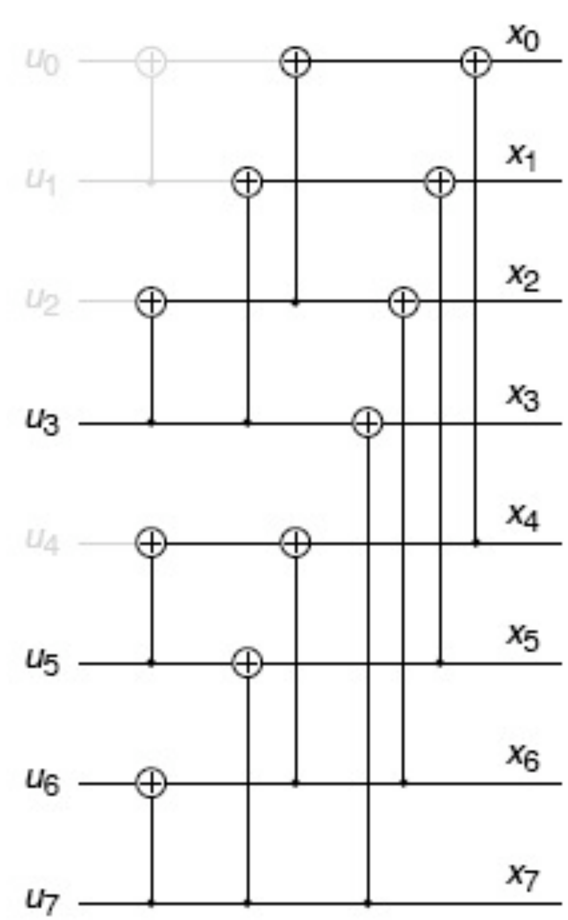
Codes from Kronecker Product of G_2

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\bar{x} = [0\ 0\ 0\ u_3\ 0\ u_5\ u_6\ u_7] G_2^{\otimes 3}$$



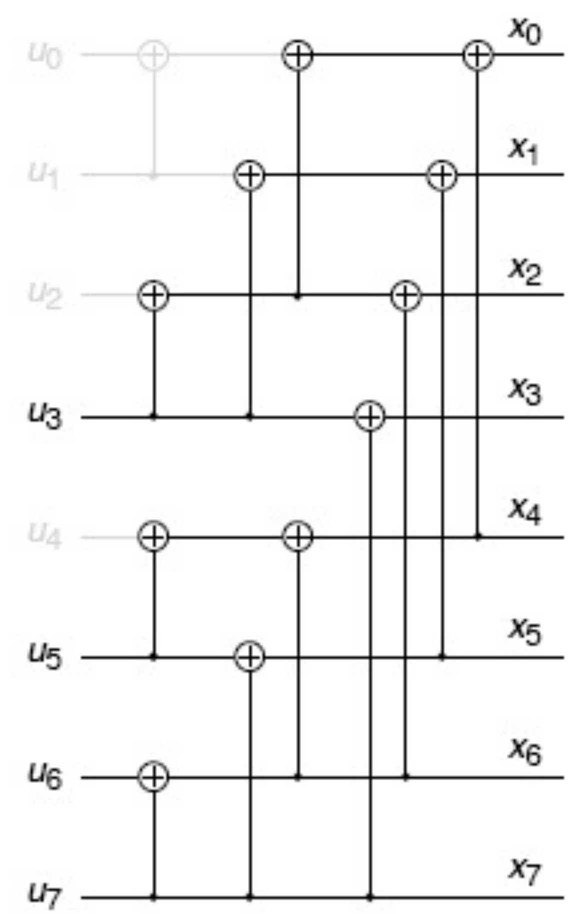
Codes from Kronecker Product of G_2

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\bar{x} = [0\ 0\ 0\ u_3\ 0\ u_5\ u_6\ u_7] G_2^{\otimes 3}$$



How to choose the rows?

Reed-Muller Codes

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

How to choose the rows?

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\bar{x} = [u_0 \; u_1 \; u_2 \; u_3 \; u_4 \; u_5 \; u_6 \; u_7] G_2^{\otimes 3}$$

Reed-Muller Codes

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

How to choose the rows?

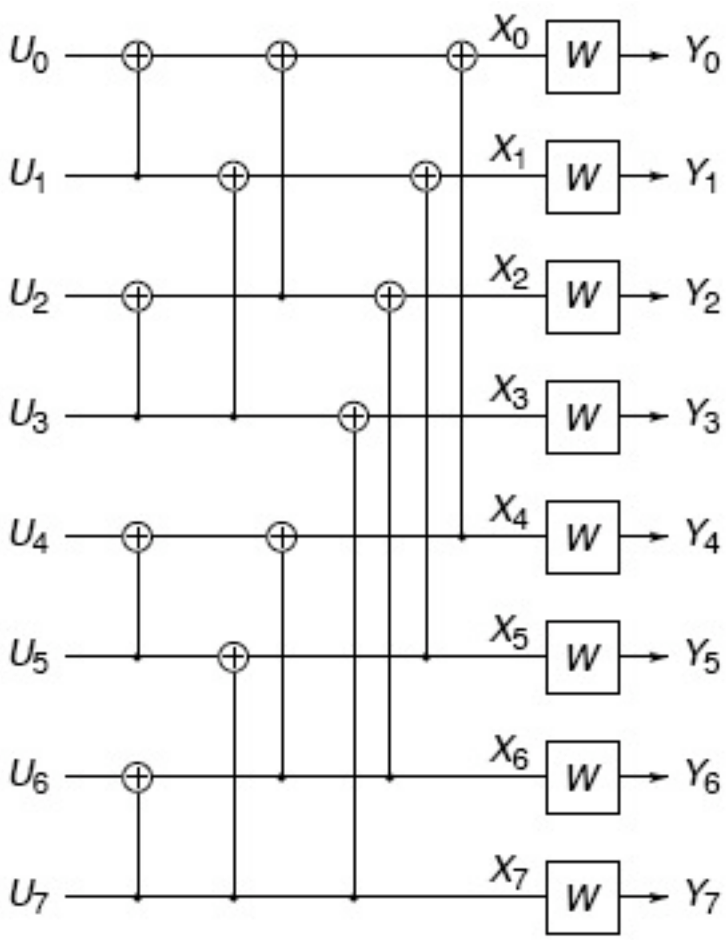
$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

choose rows of largest weight

$$\bar{x} = [0\ 0\ 0\ u_3\ 0\ u_5\ u_6\ u_7] G_2^{\otimes 3}$$

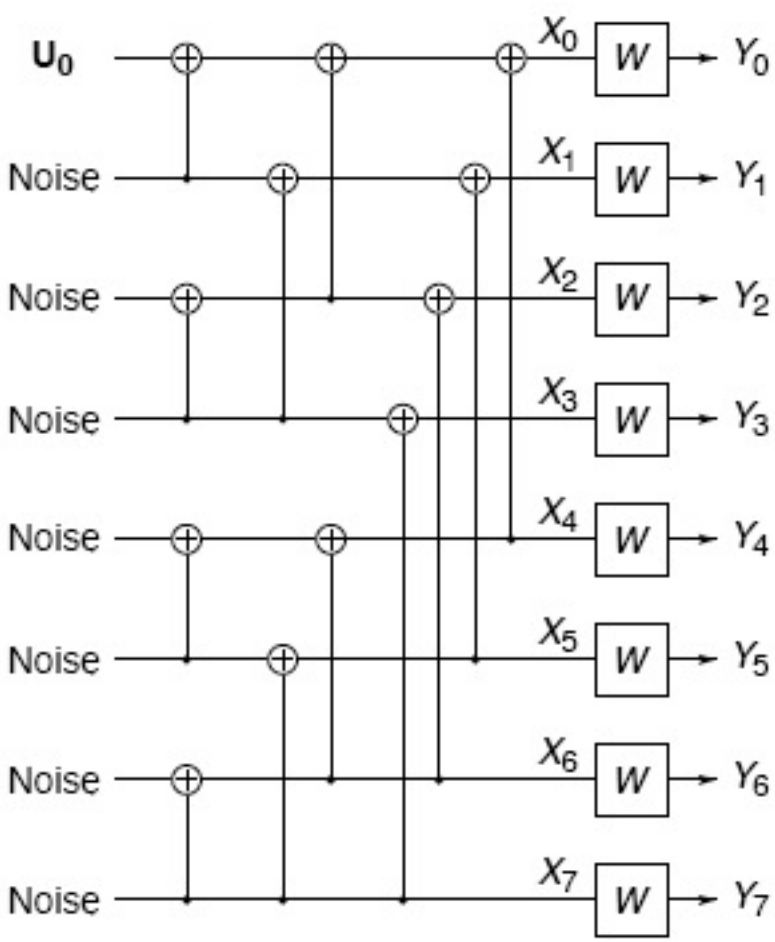
Polar Codes

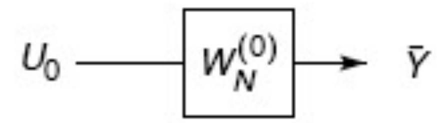
Polar Codes



W -- BMS channel

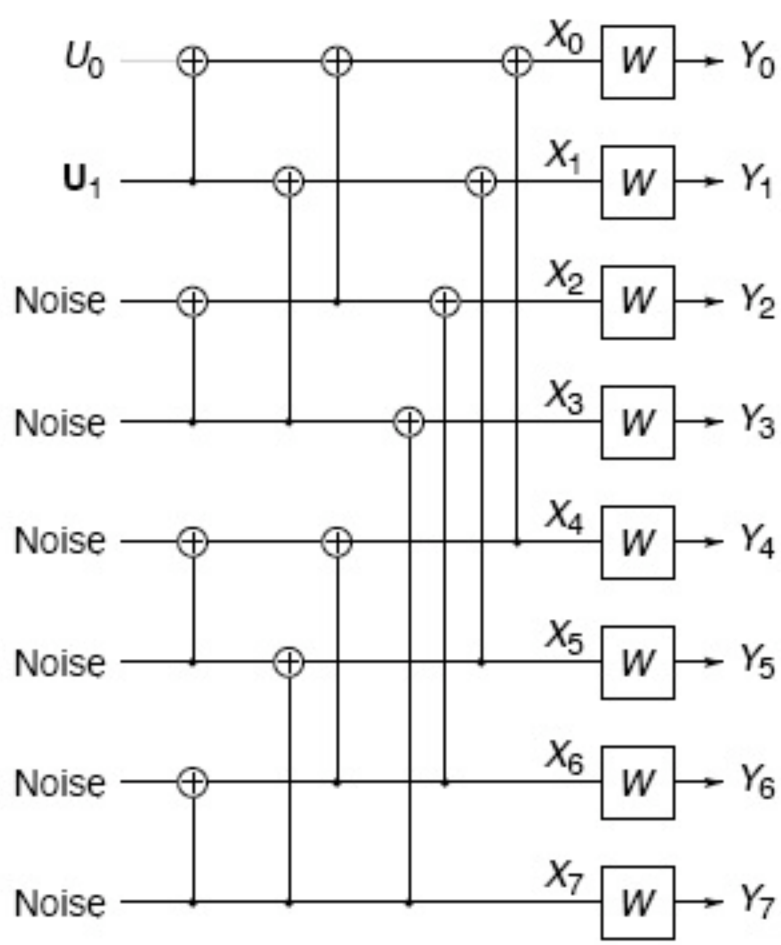
Polar Codes

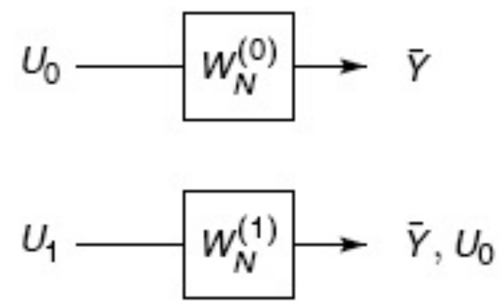




W -- BMS channel

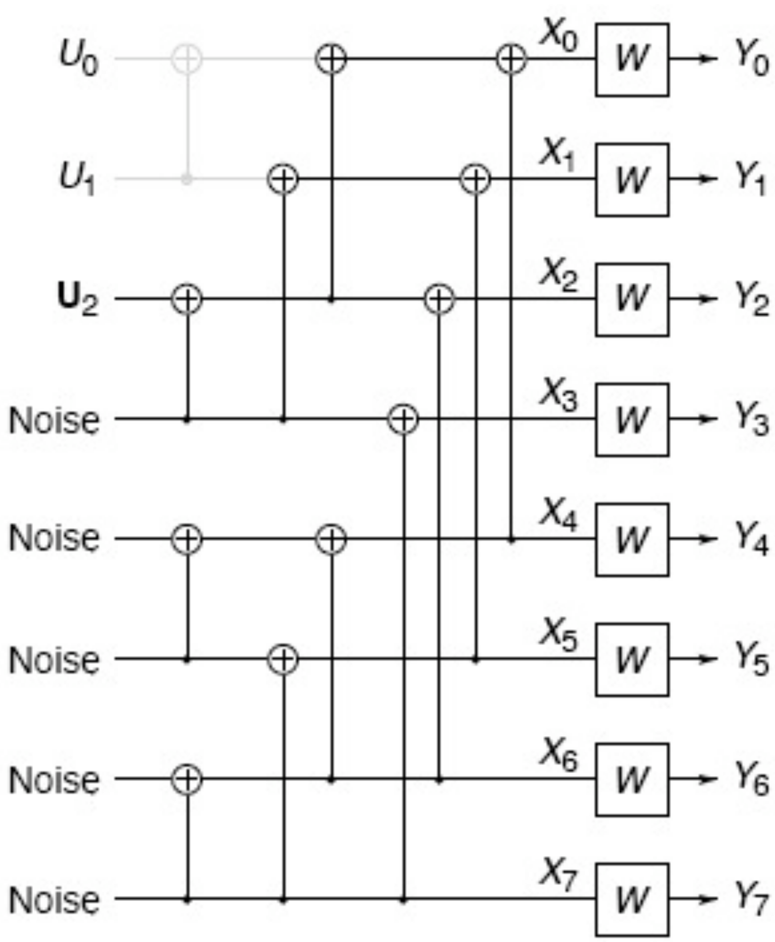
Polar Codes

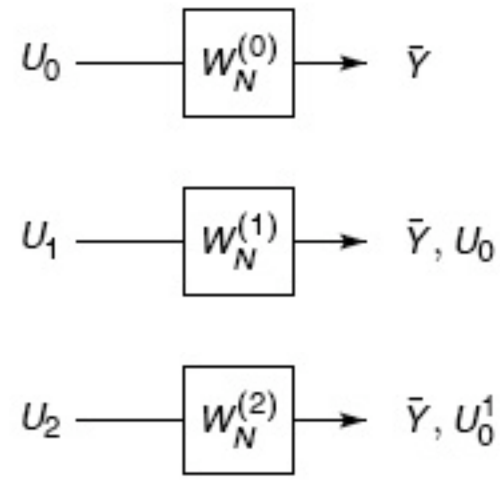




W -- BMS channel

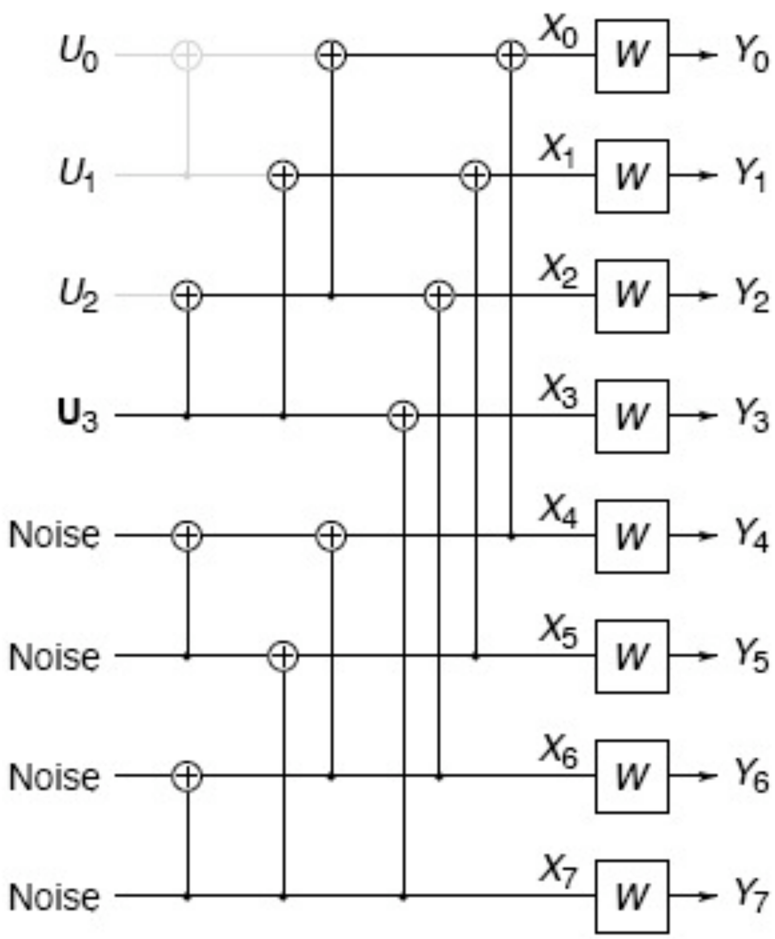
Polar Codes

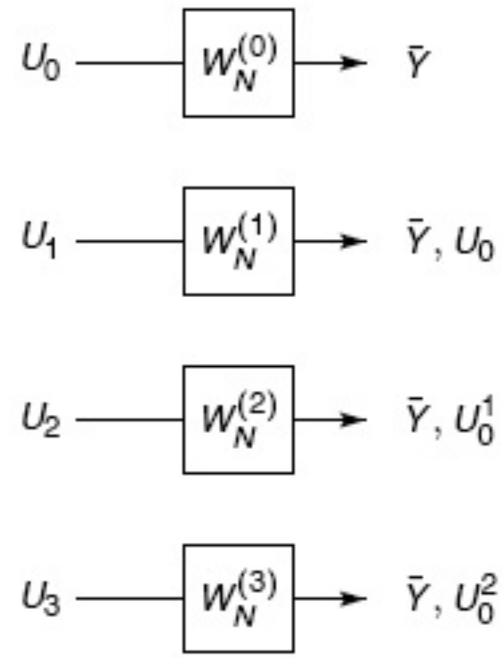




W -- BMS channel

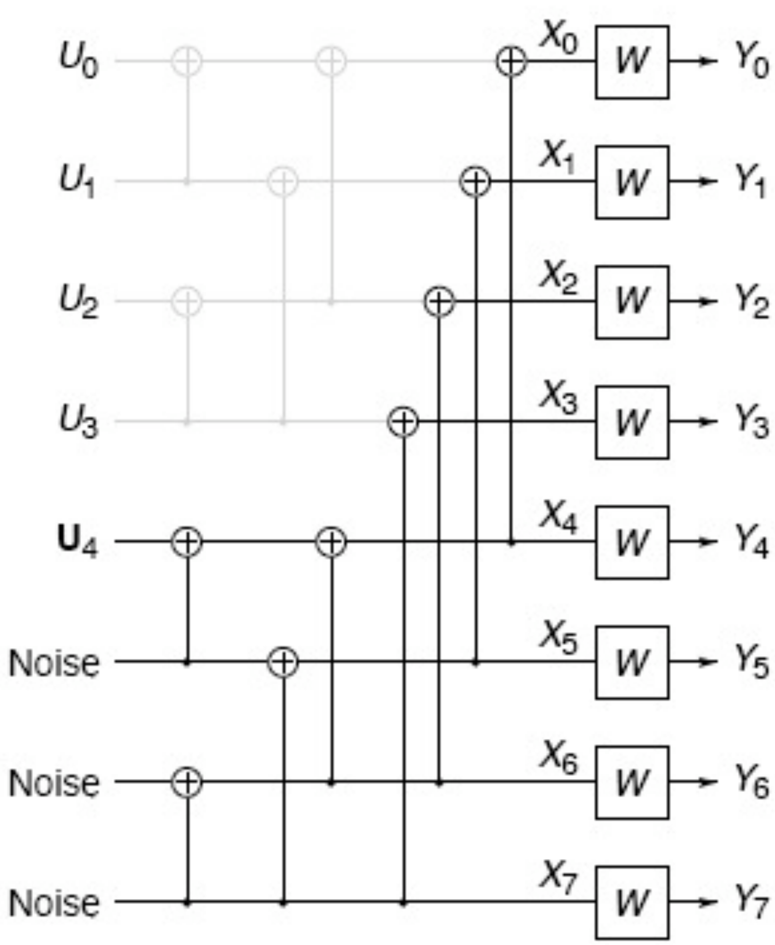
Polar Codes



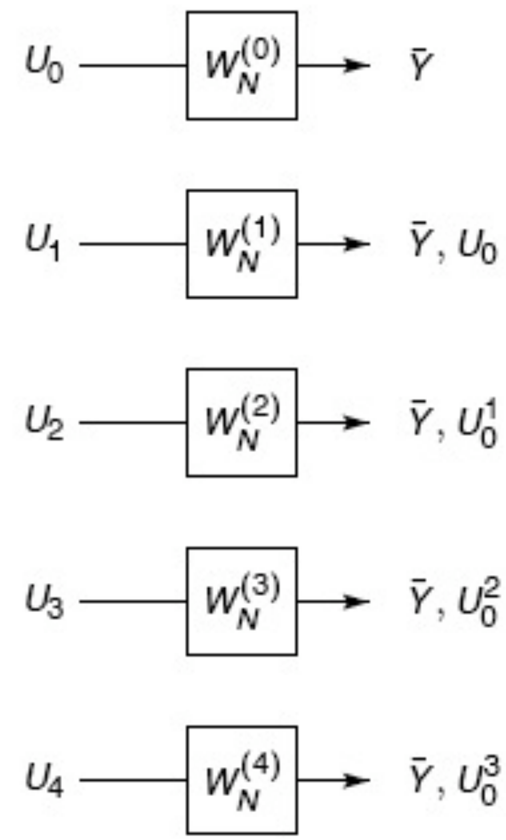


W -- BMS channel

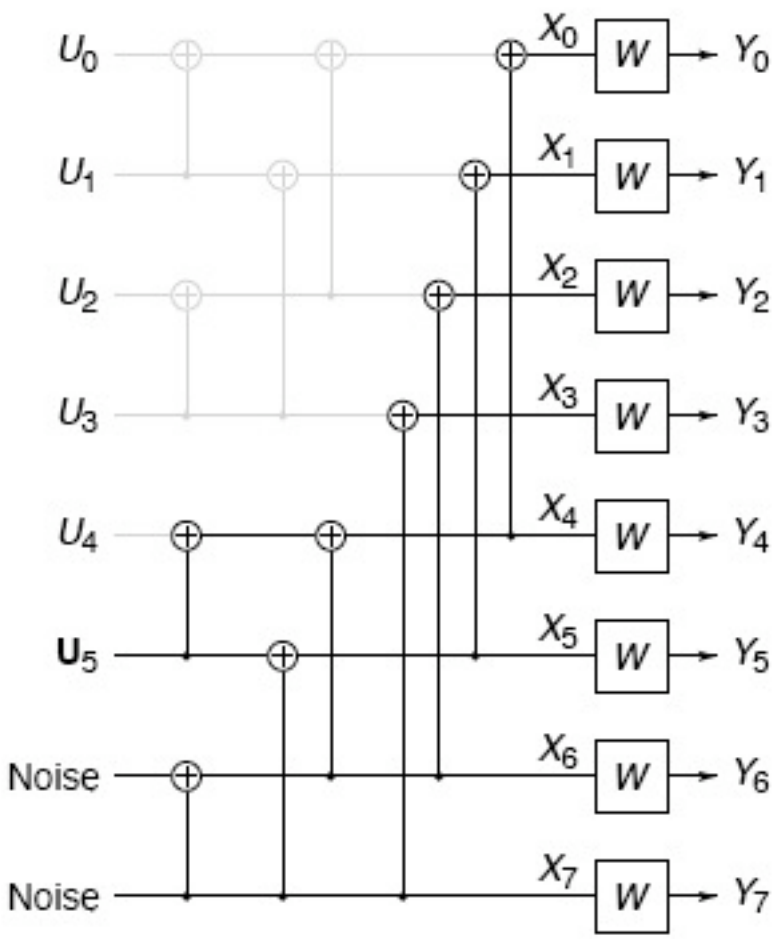
Polar Codes

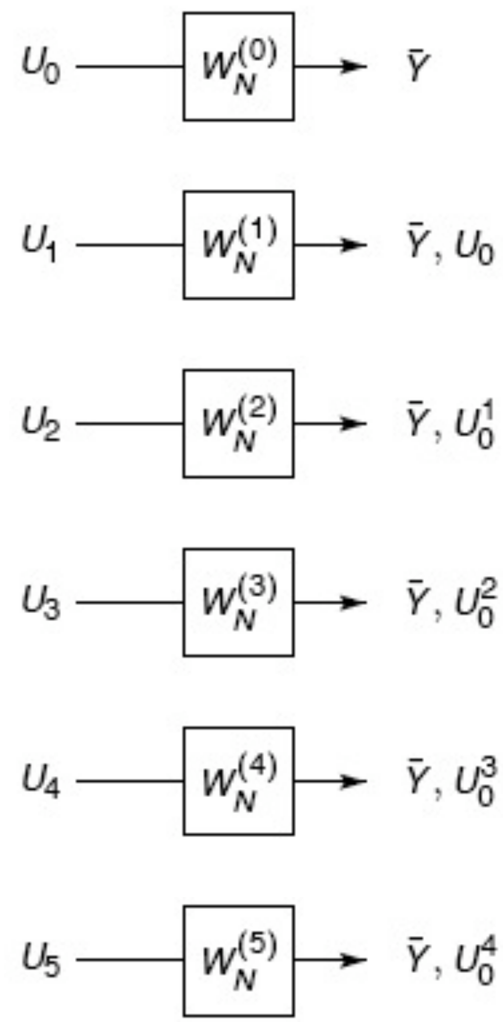


W -- BMS channel



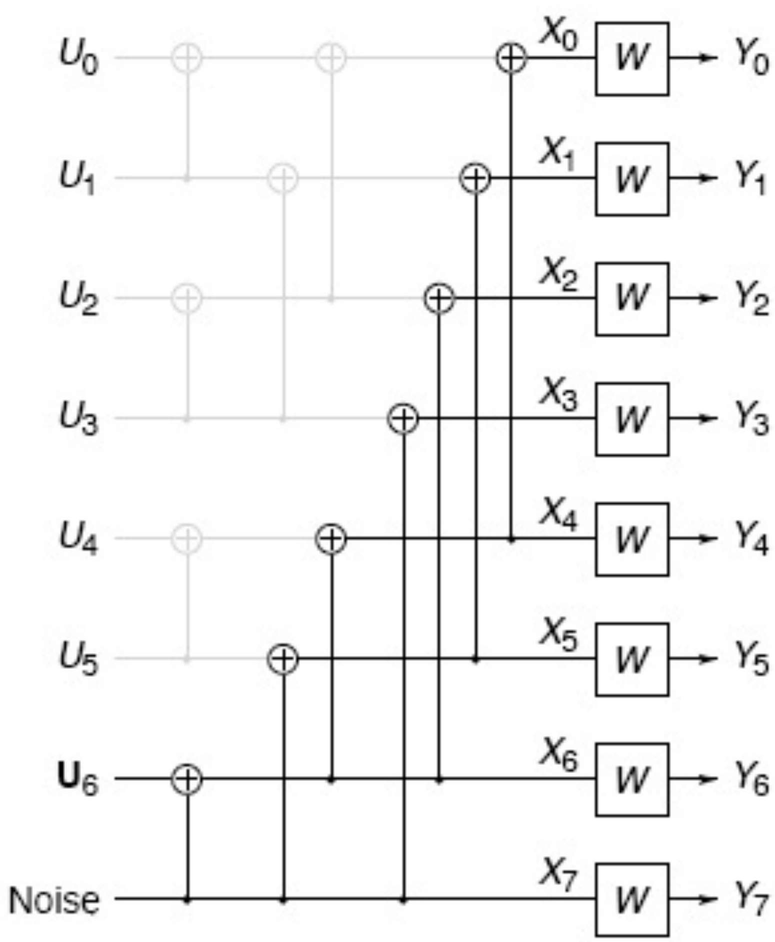
Polar Codes



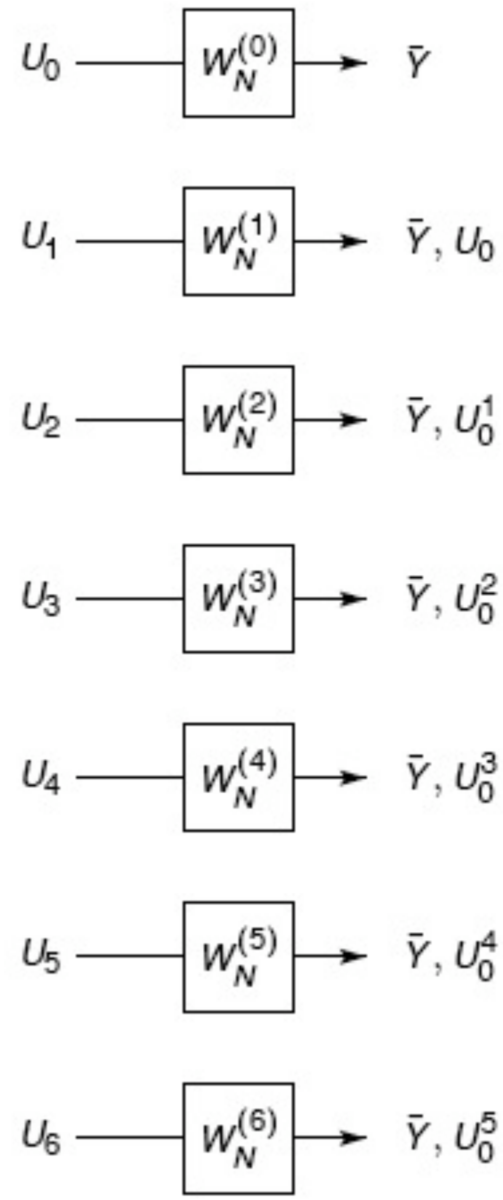


W -- BMS channel

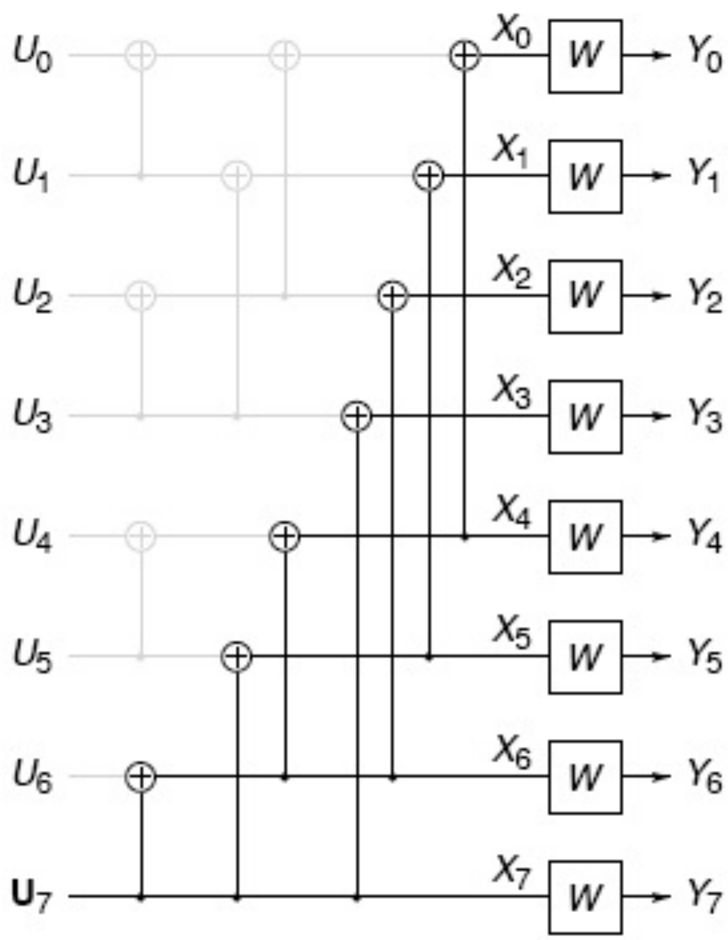
Polar Codes



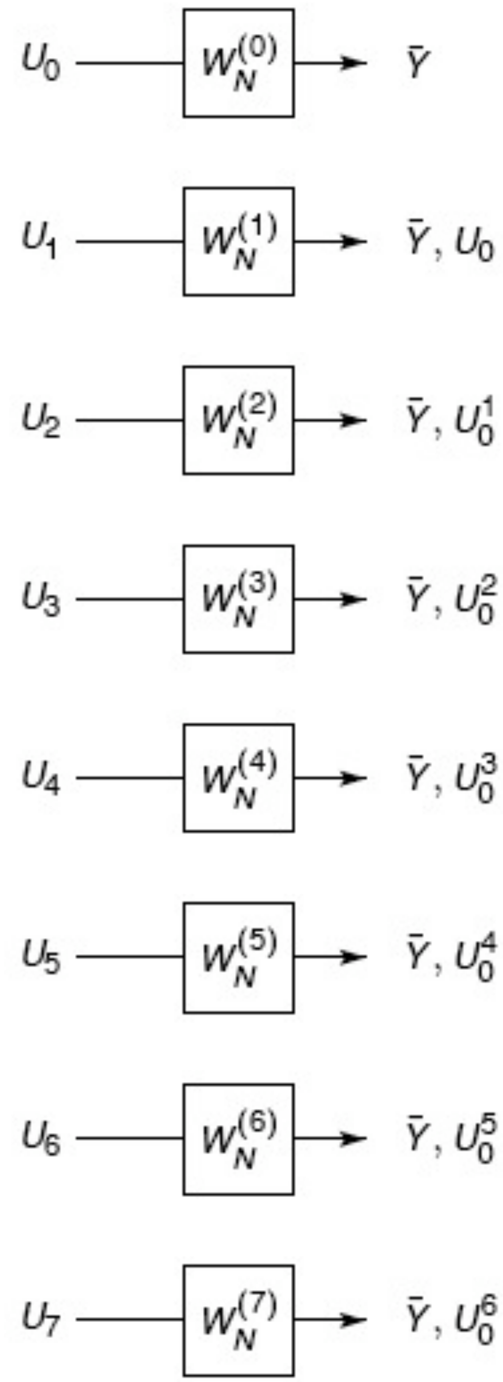
W -- BMS channel



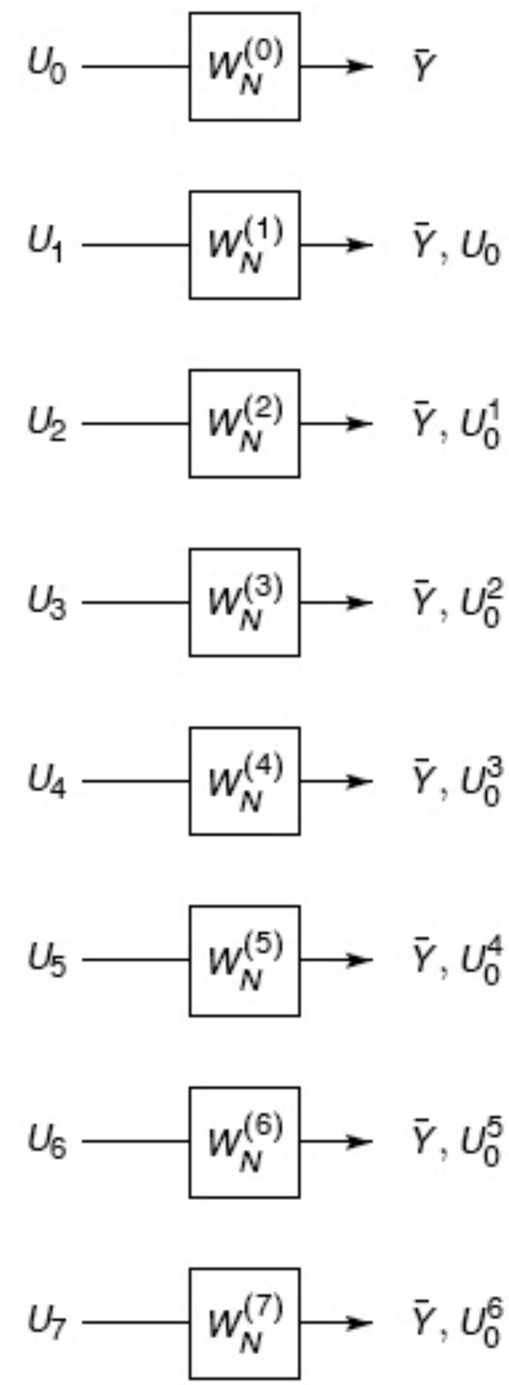
Polar Codes



W -- BMS channel

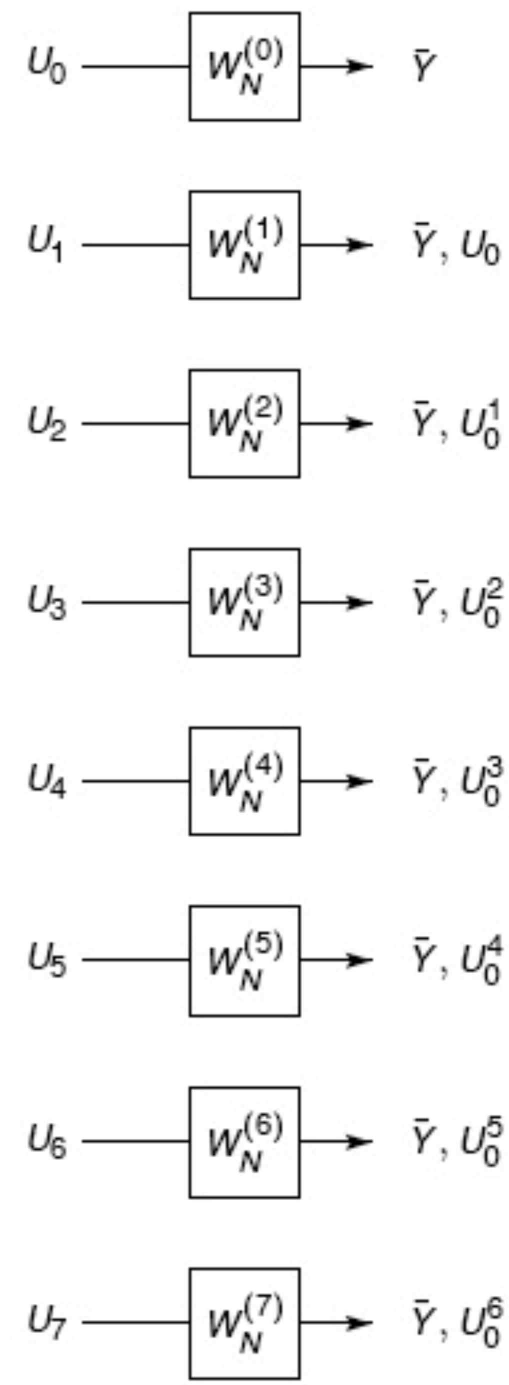


Channel Polarization



Channel Polarization

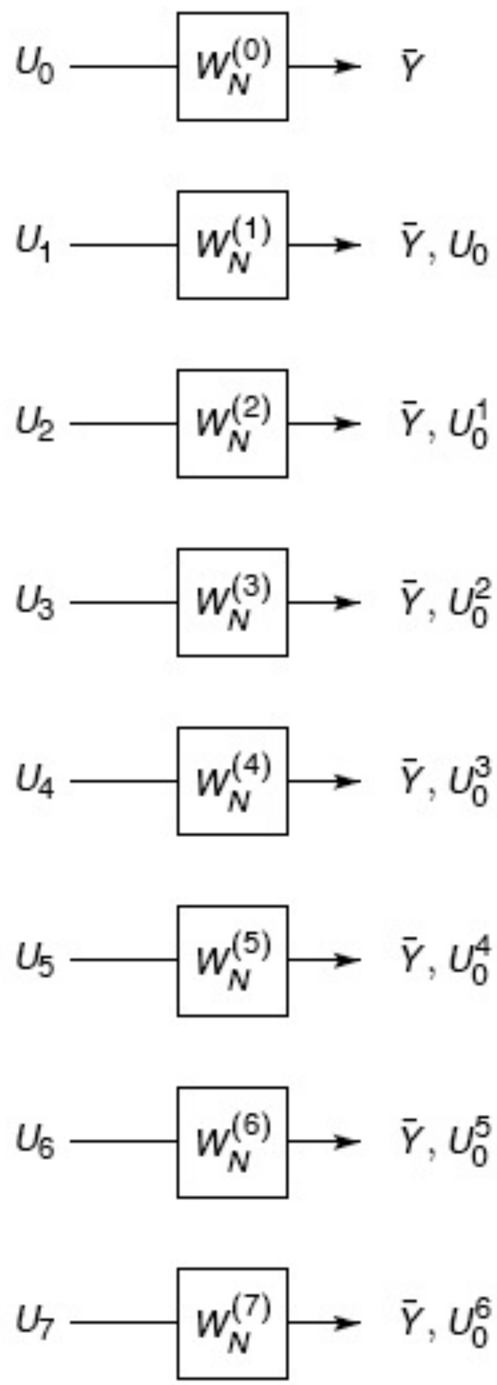
as $N \to \infty$, channels polarize, either completely noisy or noise-free



Channel Polarization

as $N \to \infty$, channels polarize, either completely noisy or noise-free

fraction of good channels approaches capacity $I(W)$

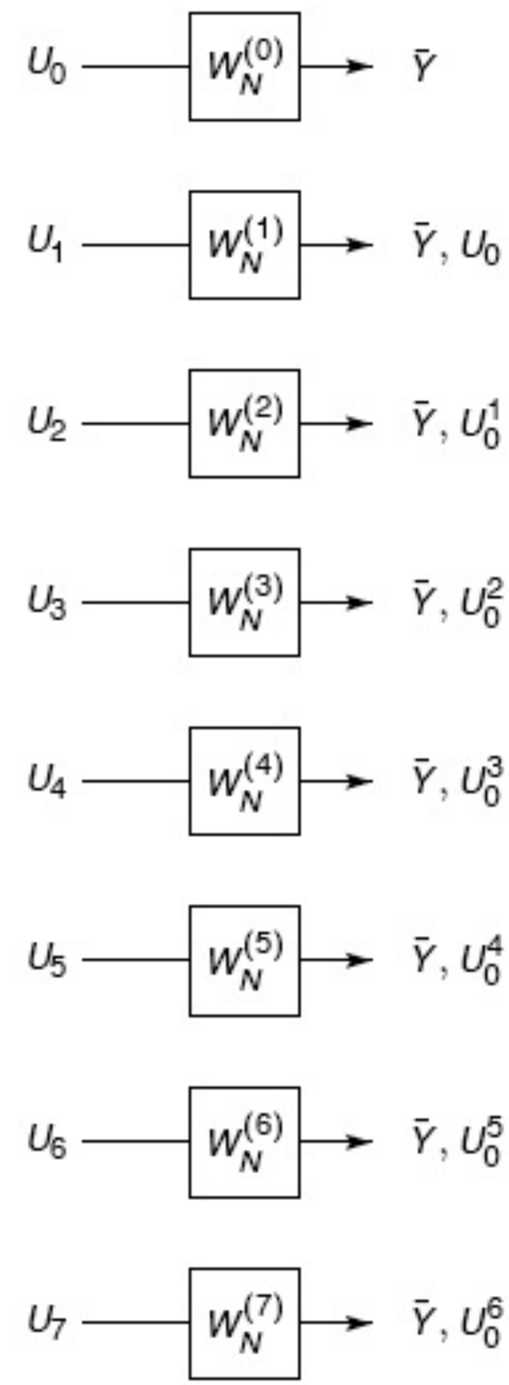


Channel Polarization

as $N \to \infty$, channels polarize, either completely noisy or noise-free

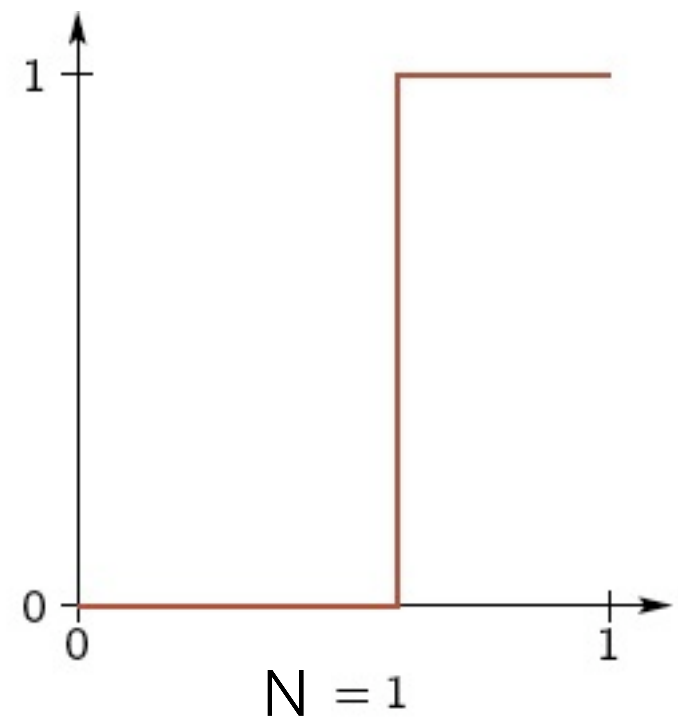
fraction of good channels approaches capacity $I(W)$

fix bad channels and transmit uncoded bits over the good ones

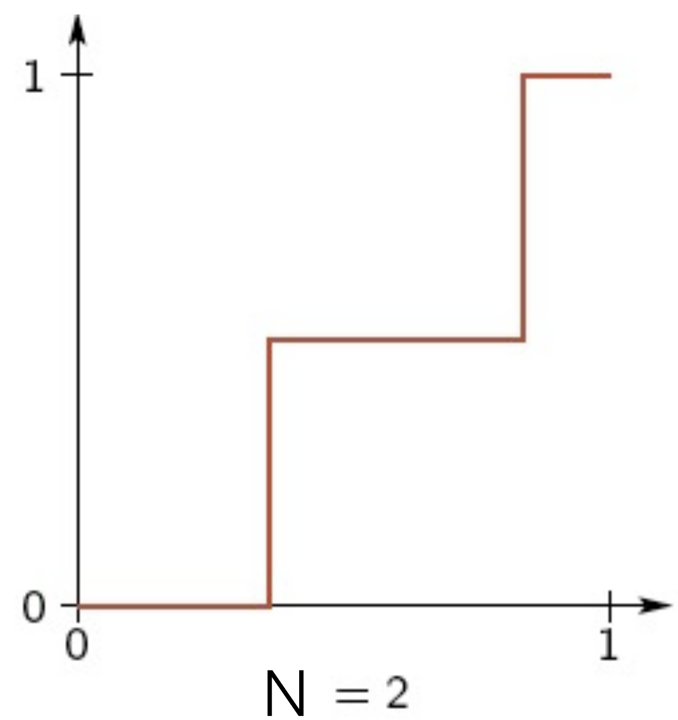


Channel Polarization

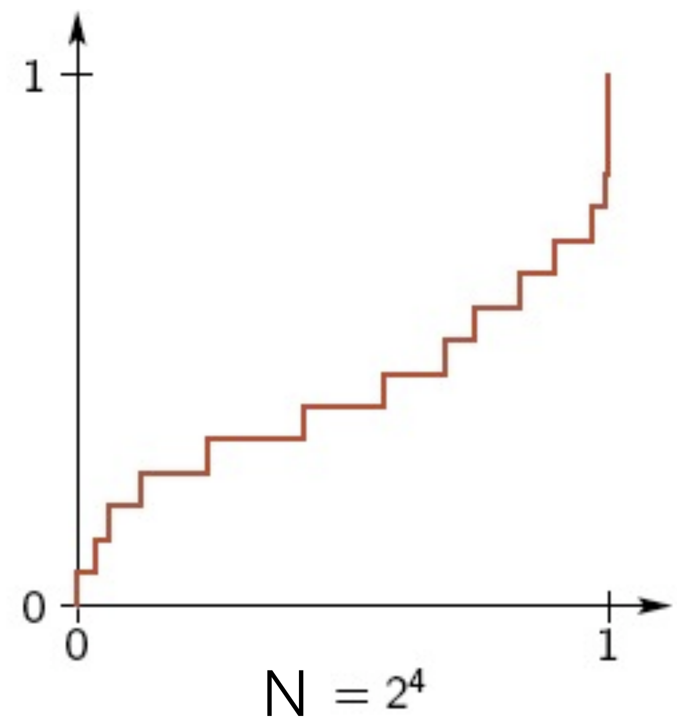
Channel Polarization



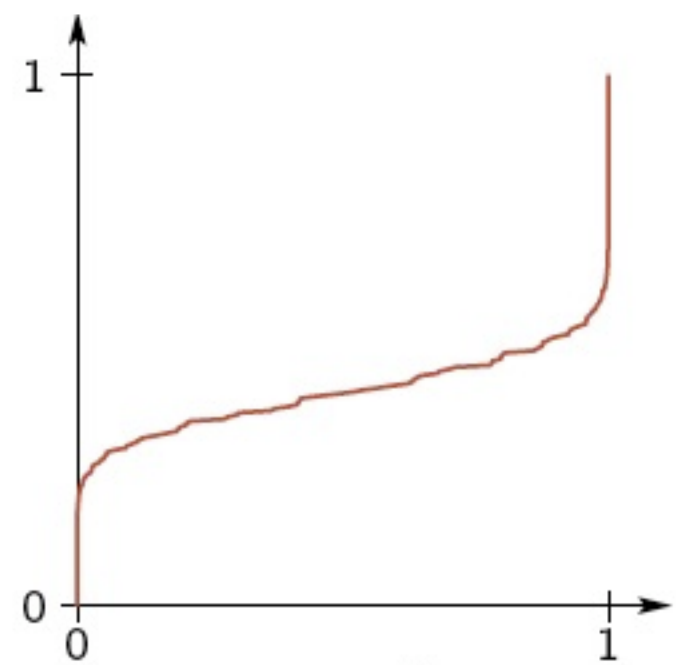
Channel Polarization



Channel Polarization

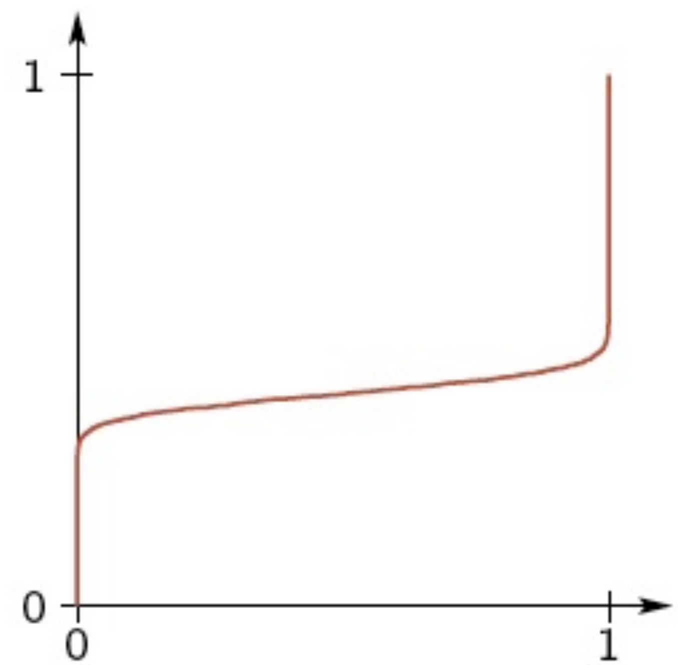


Channel Polarization



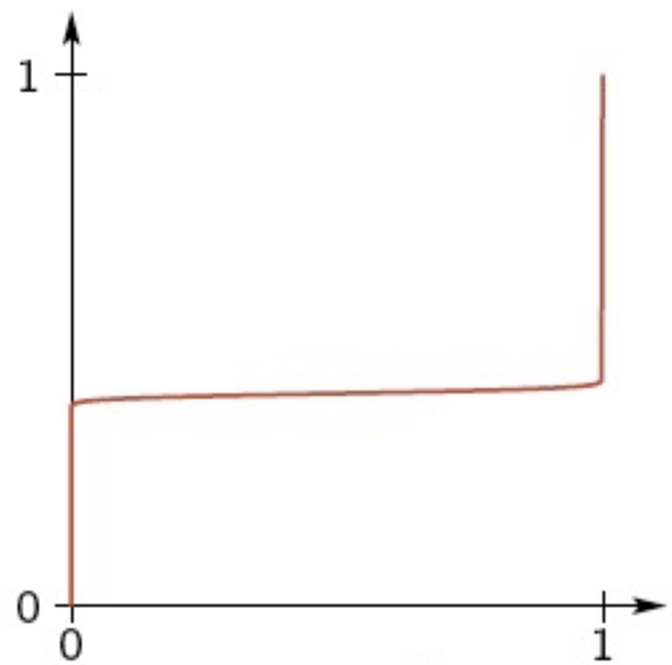
$N = 2^8$

Channel Polarization



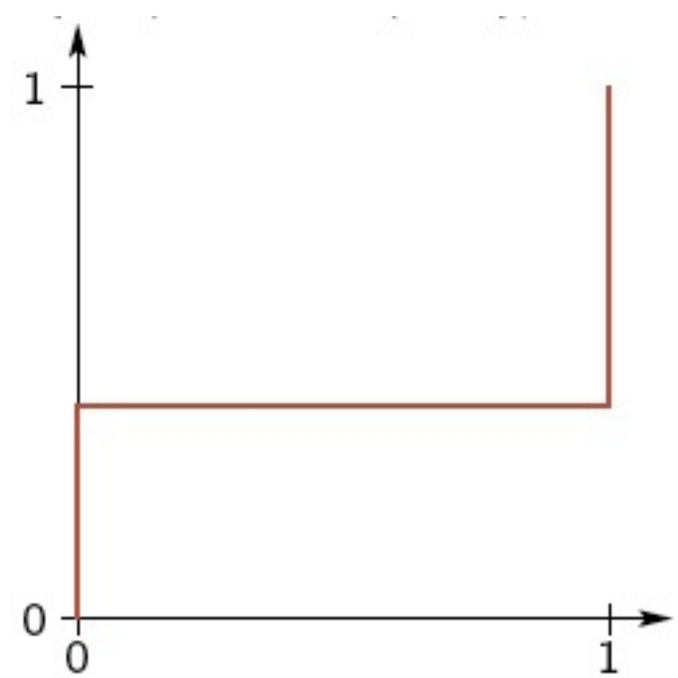
$N = 2^{12}$

Channel Polarization

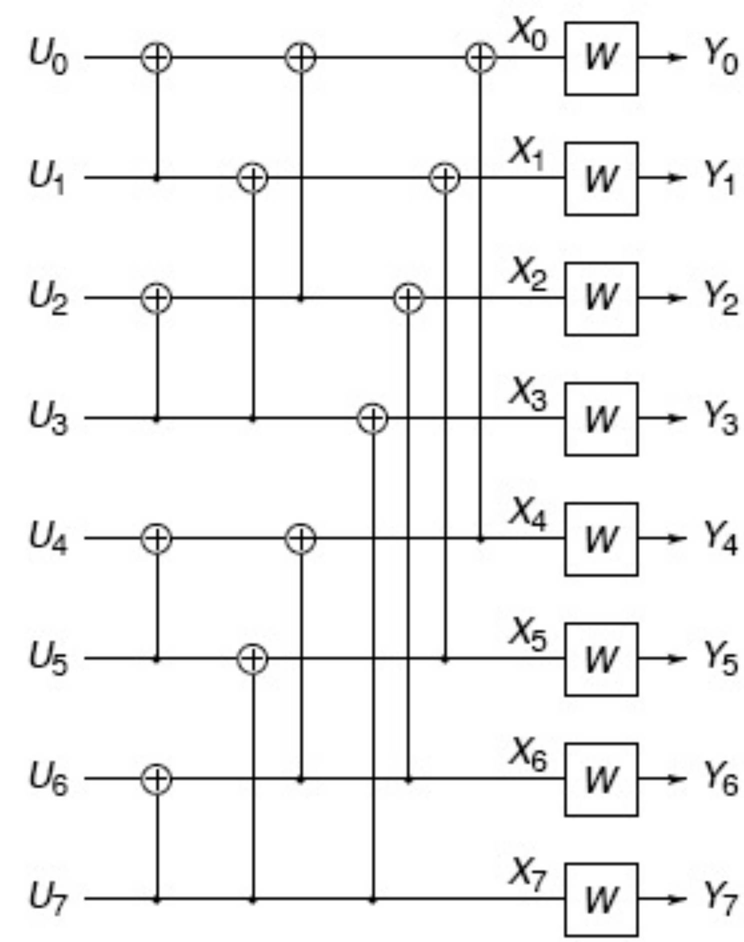


$N = 2^{20}$

Channel Polarization

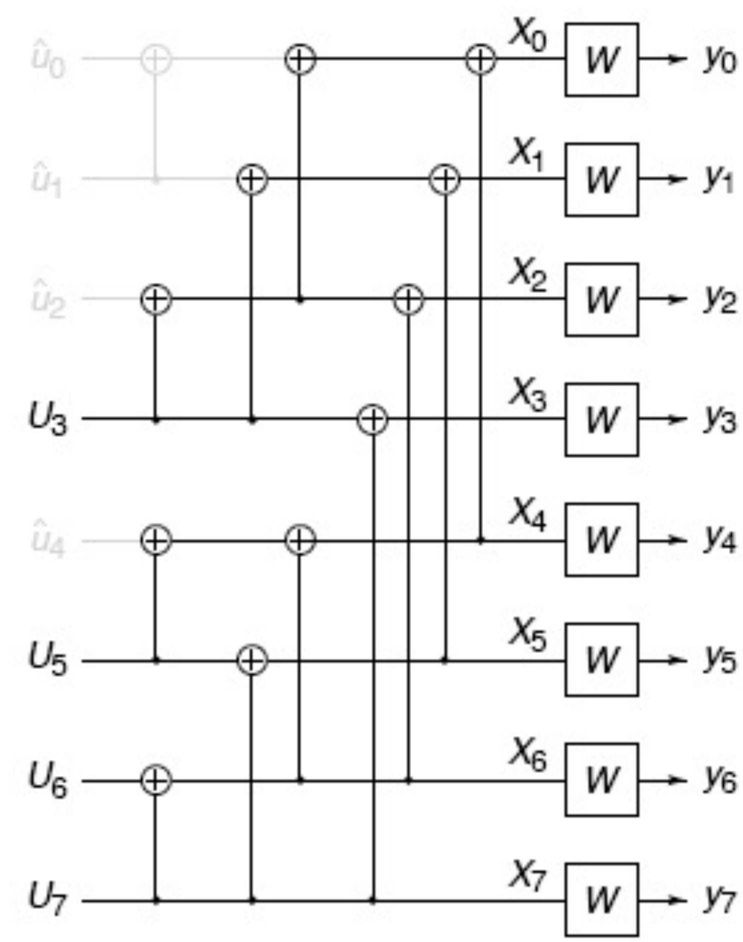


Successive Decoding



Successive Decoding

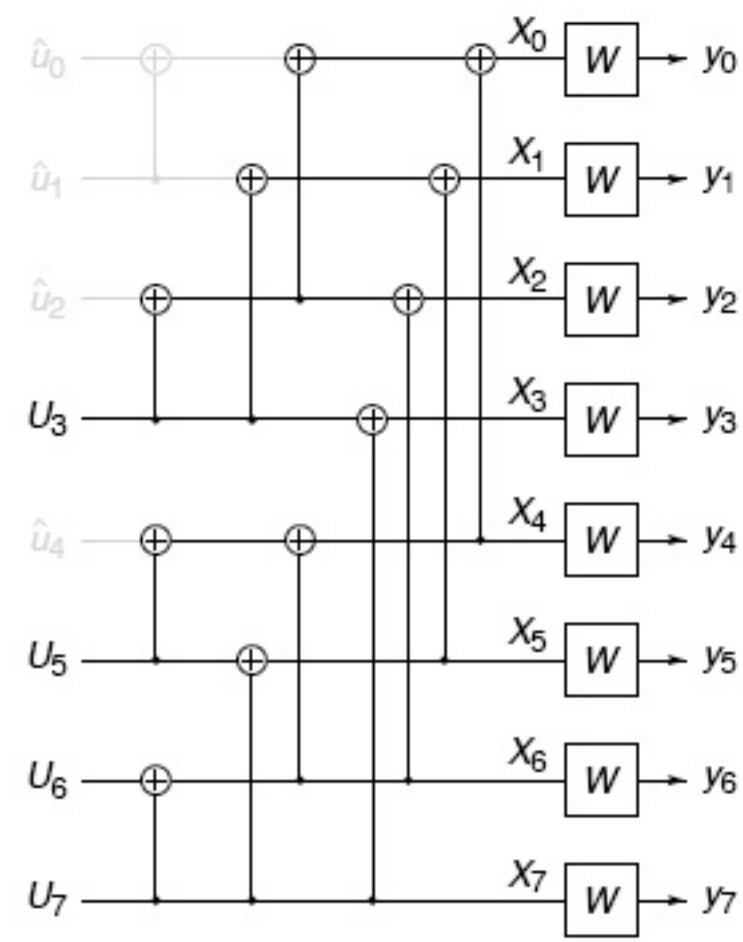
$$F = \{0, 1, 2, 4\}$$



Successive Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$



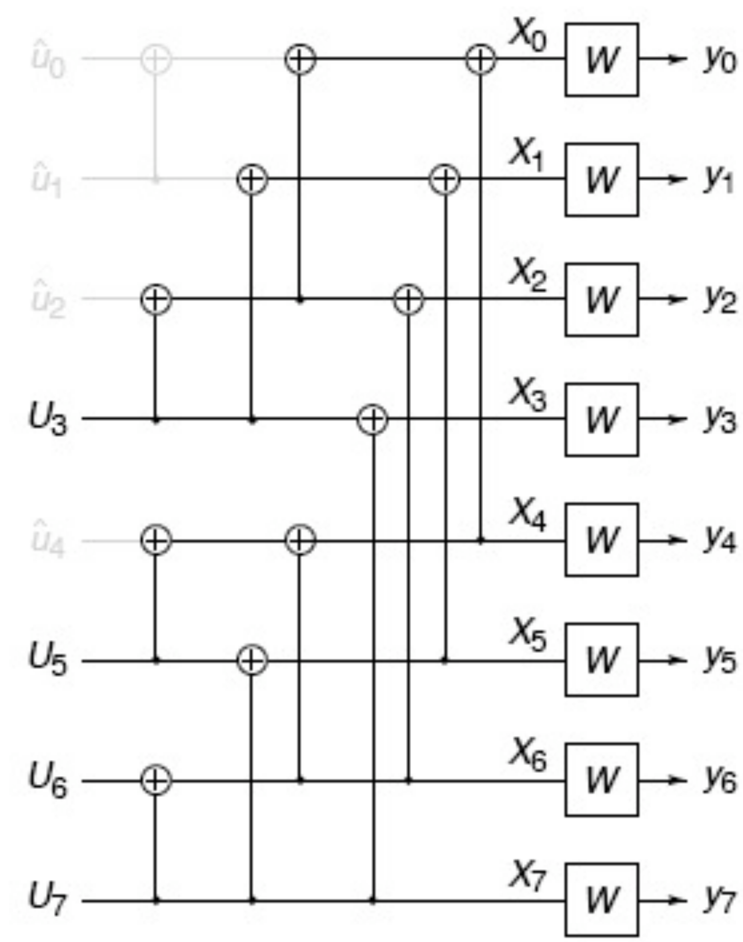
Successive Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N-1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1},\bar{y})}{P(1|\hat{u}_0^{i-1},\bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



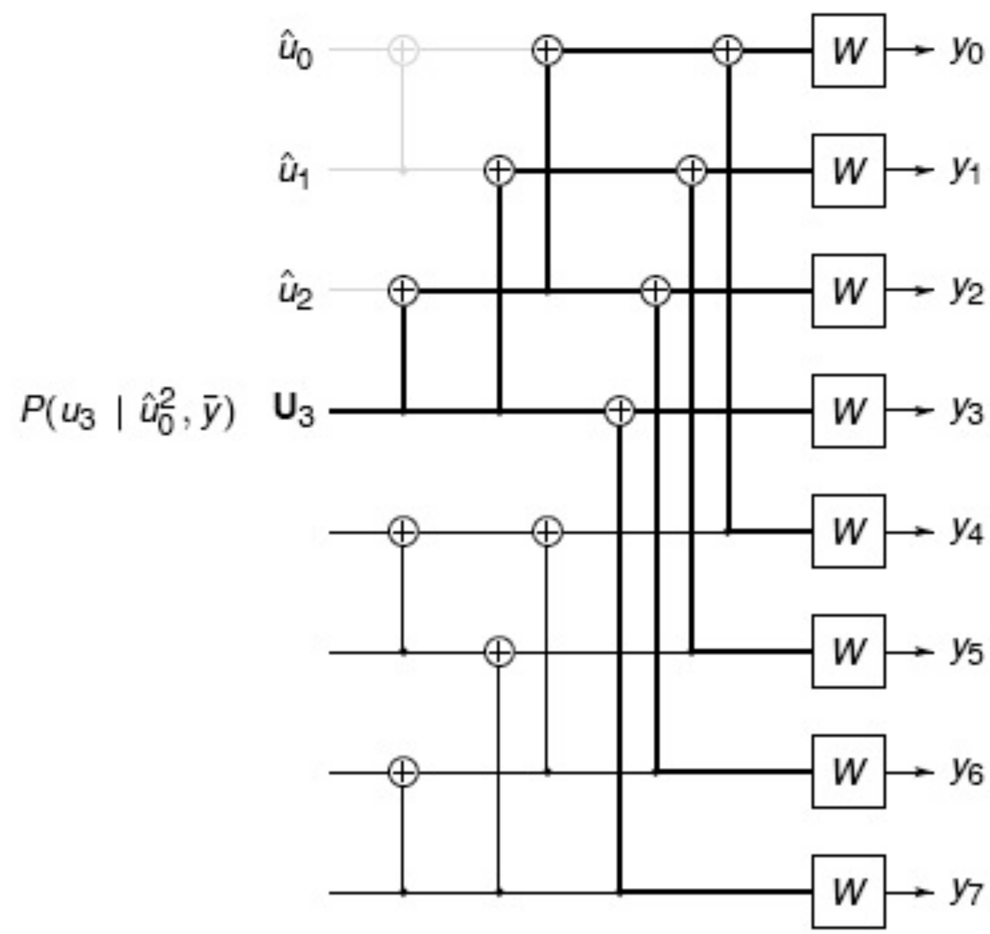
Successive Decoding

$$F = \{0, 1, 2, 4\}$$

From 0 till $N-1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



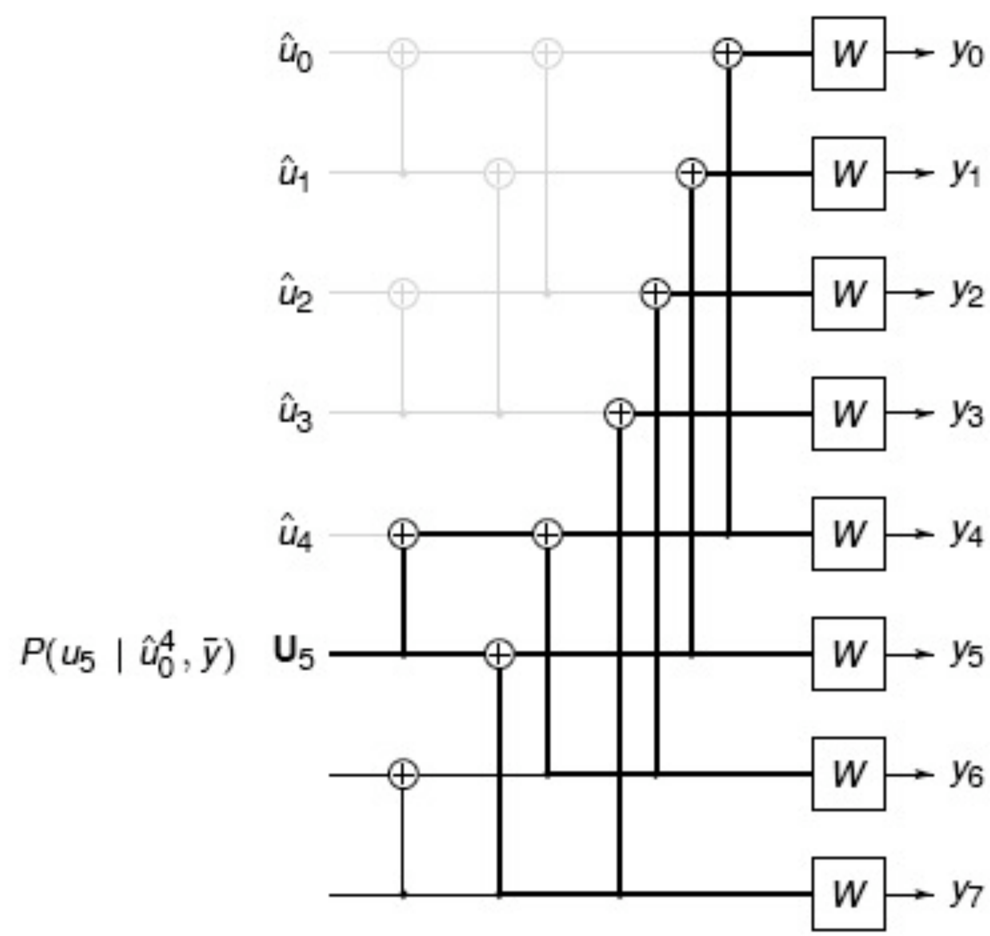
Successive Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



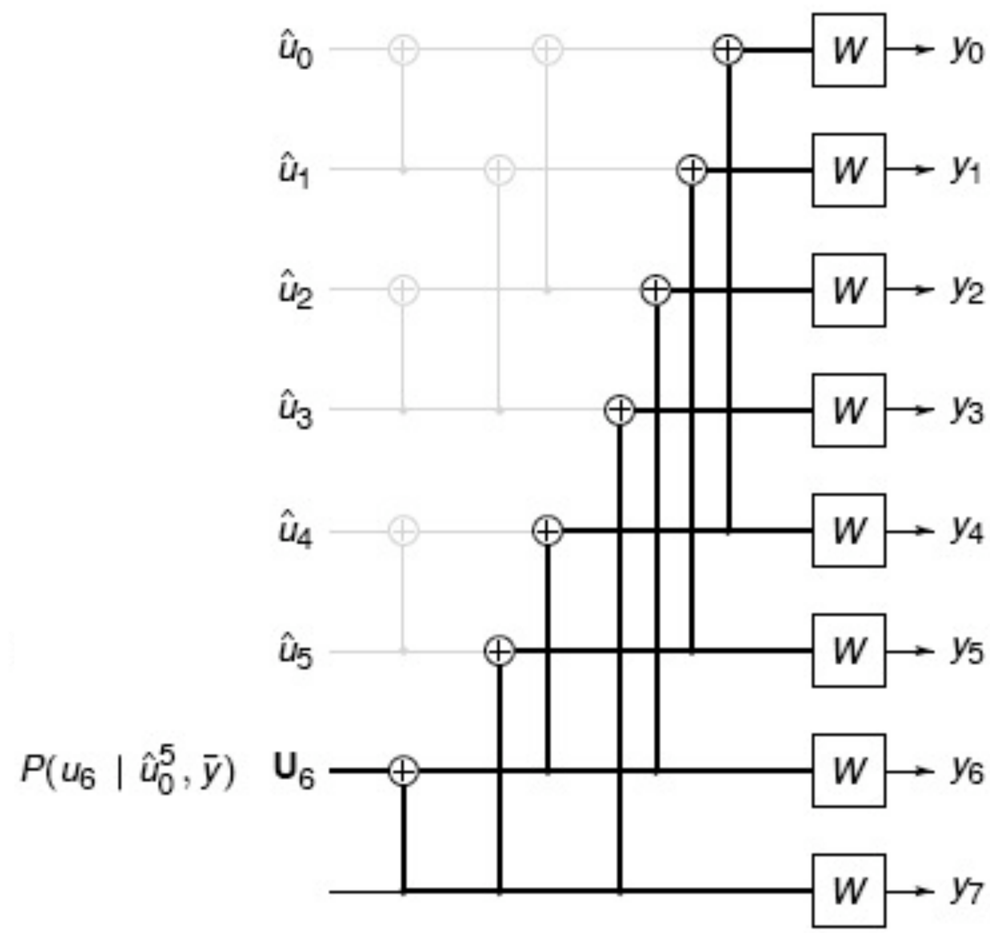
Successive Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N-1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1},\bar{y})}{P(1|\hat{u}_0^{i-1},\bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



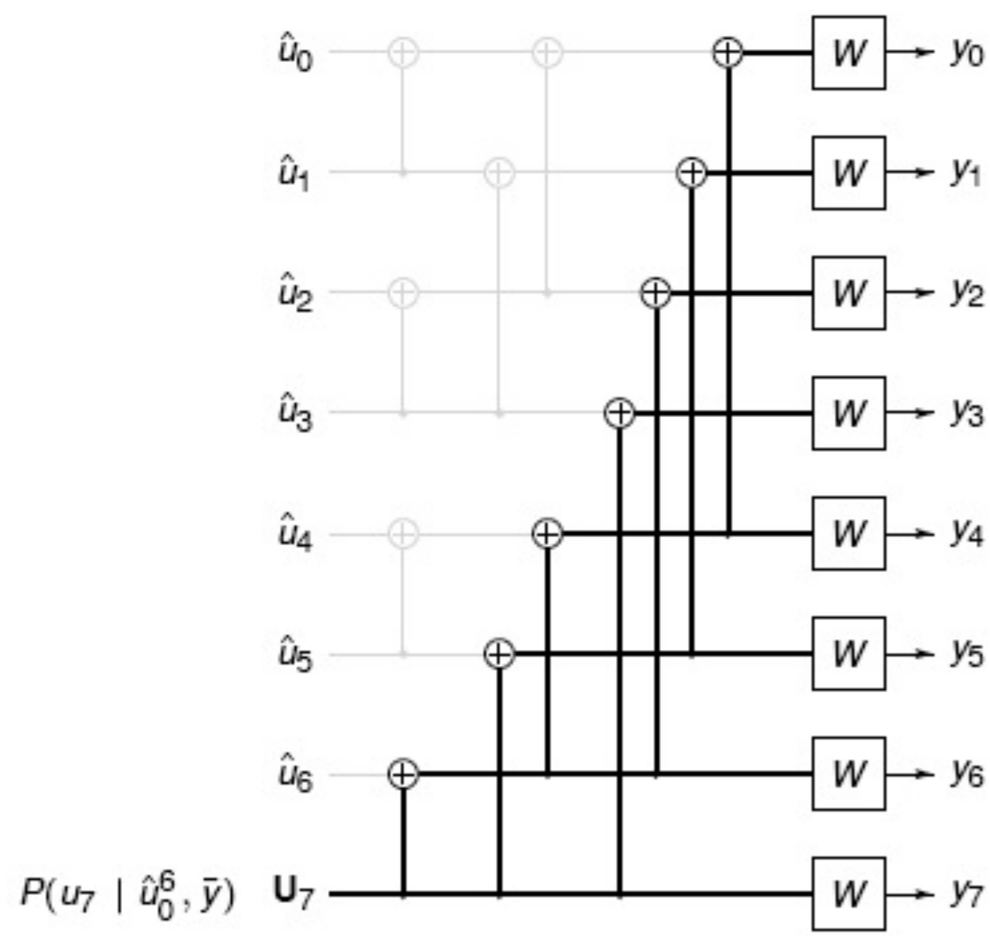
Successive Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



Successive Decoding

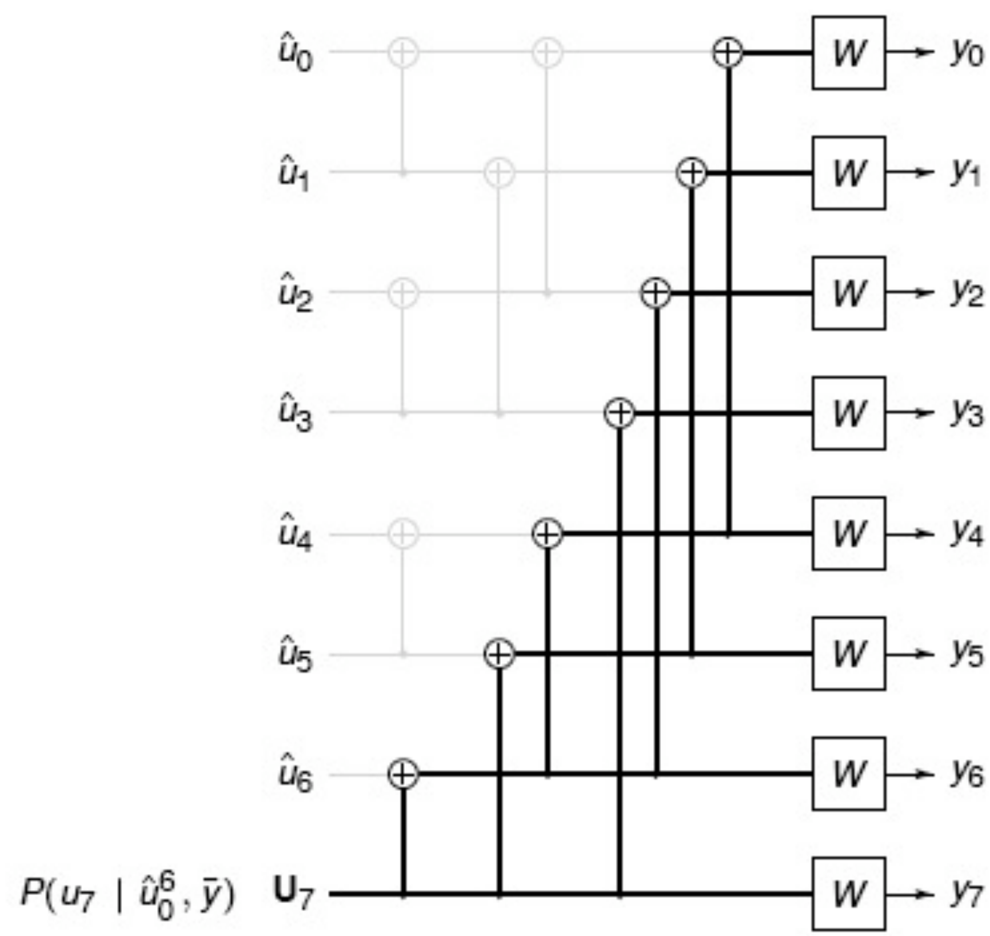
$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$

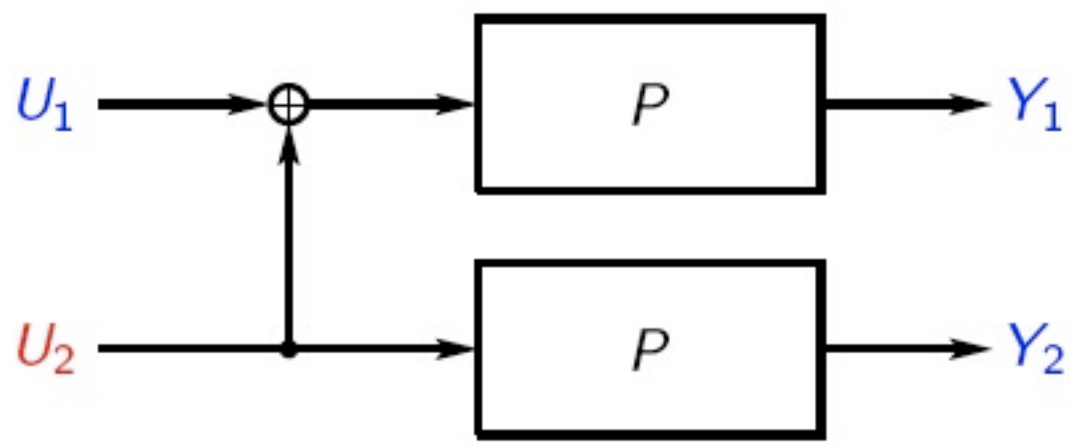
if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$

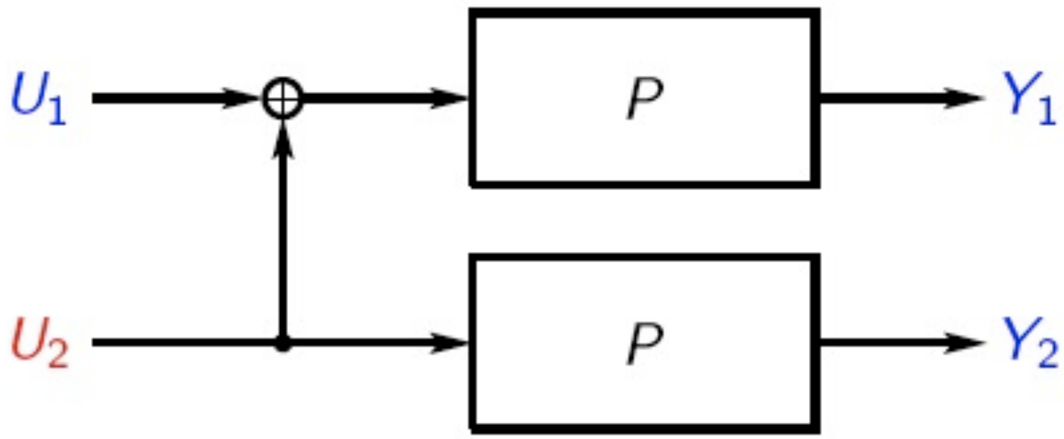
Complexity $O(N \log N)$



More on Polarization



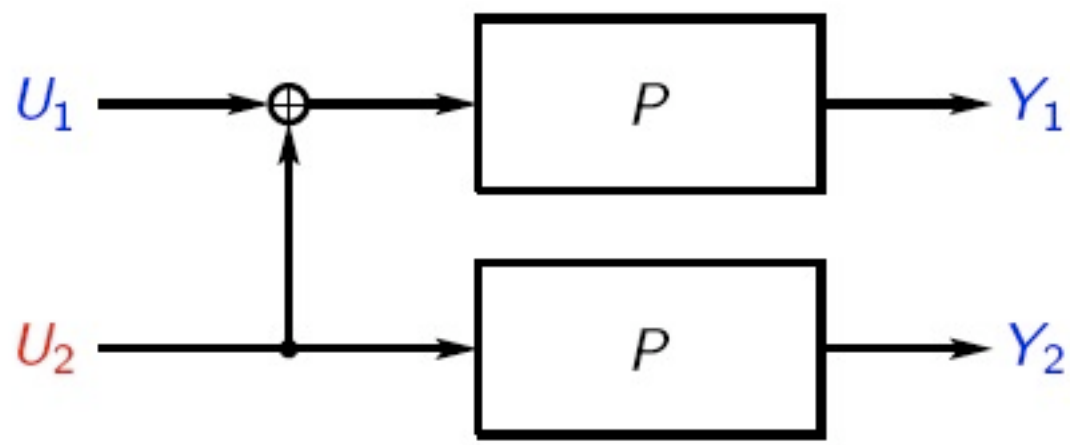
More on Polarization



$$P^-(y_1 y_2 | u_1) = \sum_{u_2} \tfrac{1}{2} P(y_1 | u_1 + u_2) P(y_2 | u_2)$$

$$P^+(y_1 y_2 u_1 | u_2) = \tfrac{1}{2} P(y_1 | u_1 + u_2) P(y_2 | u_2)$$

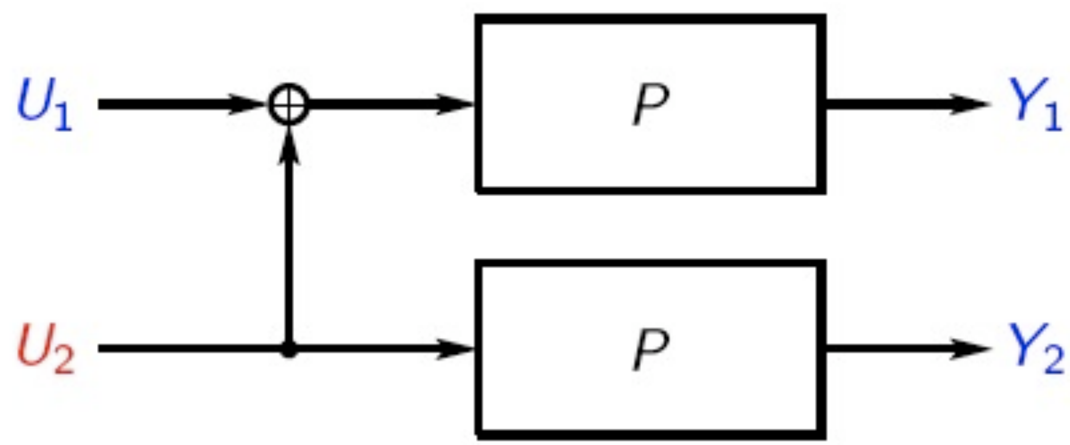
More on Polarization



$$I(P^-) = I(U_1; Y_1 Y_2)$$

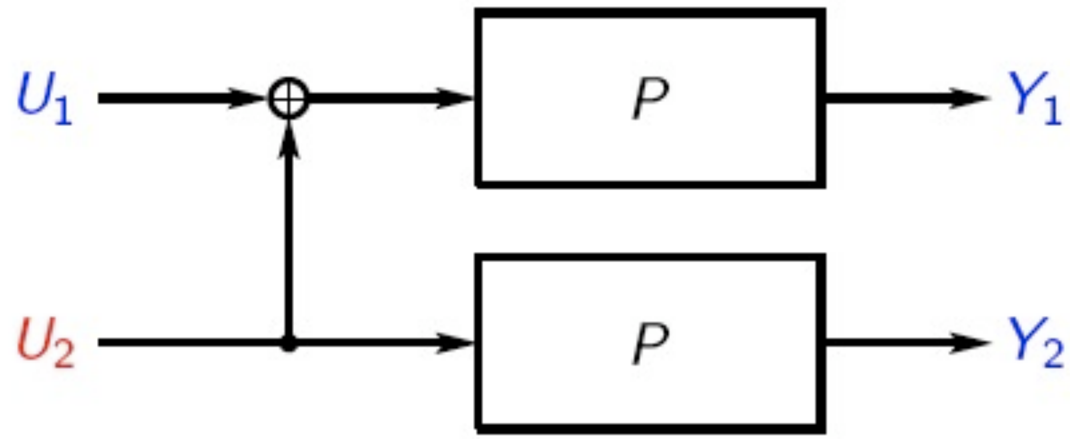
$$I(P^+) = I(U_2; Y_1 Y_2 U_1)$$

More on Polarization



$$
\begin{aligned}
I(P^-) + I(P^+) &= I(U_1 U_2; Y_1 Y_2) \\
&= I(X_1; Y_1) + I(X_2; Y_2) \\
&= 2I(P),
\end{aligned}
$$

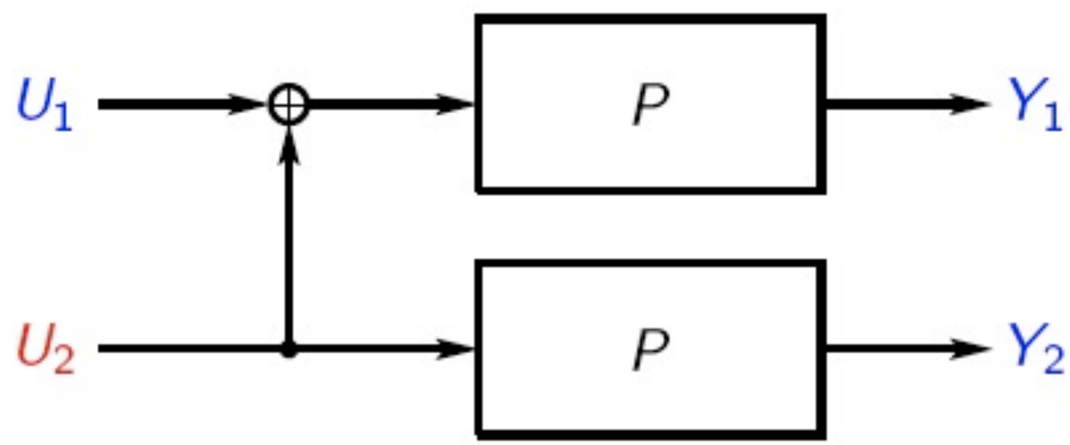
More on Polarization



$$
\begin{aligned}
I(P^-) + I(P^+) &= I(U_1 U_2; Y_1 Y_2) \\
&= I(X_1; Y_1) + I(X_2; Y_2) \\
&= 2I(P),
\end{aligned}
$$

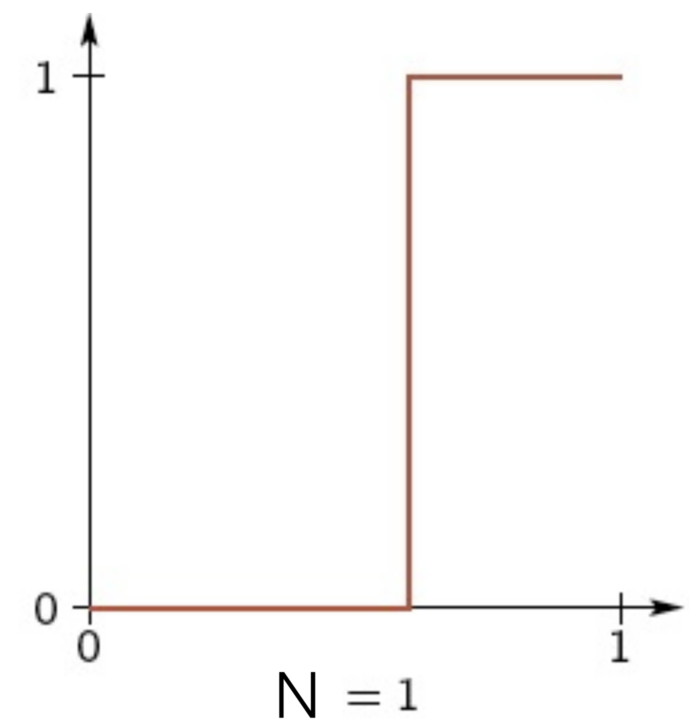
$$I(P^-) \leq I(P) \leq I(P^+)$$

More on Polarization

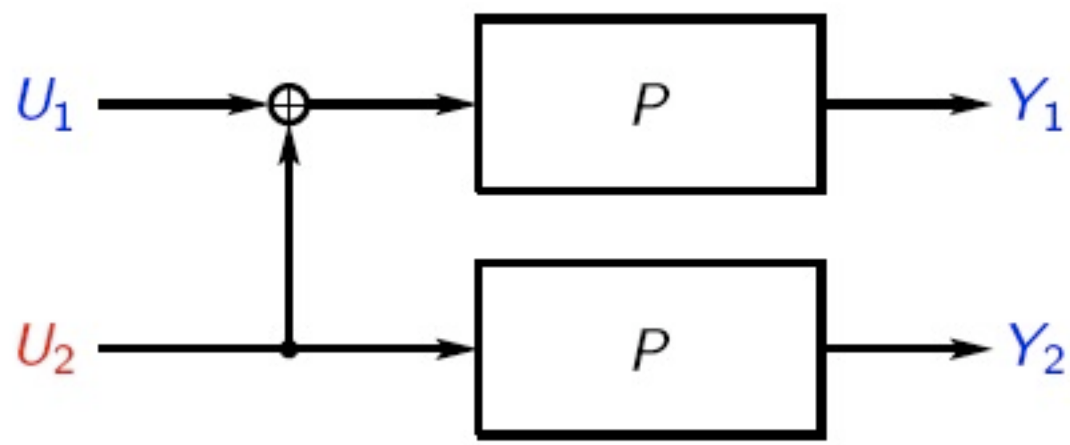


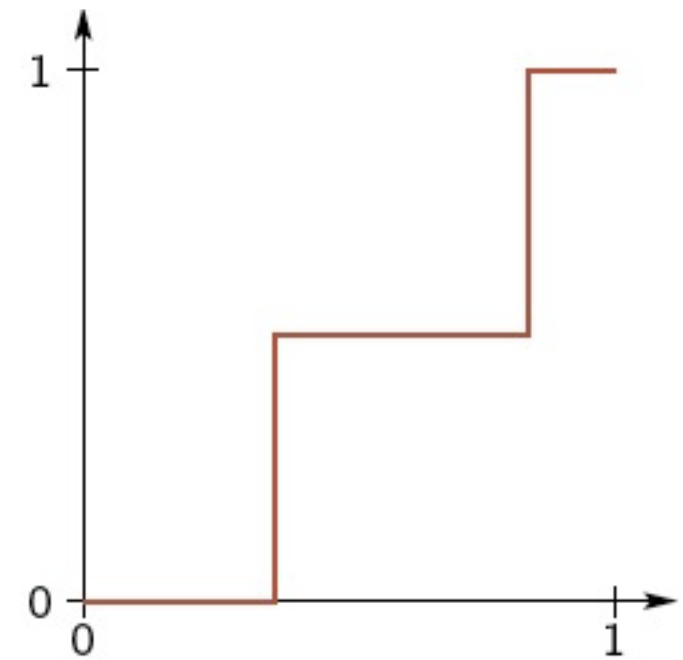
$$
\begin{aligned}
I(P^-) + I(P^+) &= I(U_1 U_2; Y_1 Y_2) \\
&= I(X_1; Y_1) + I(X_2; Y_2) \\
&= 2I(P),
\end{aligned}
$$

$$I(P^-) \leq I(P) \leq I(P^+)$$



More on Polarization





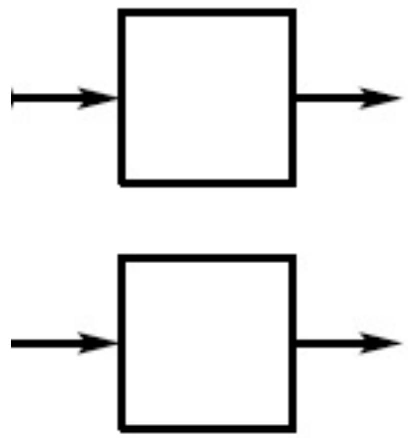
N $= 2$

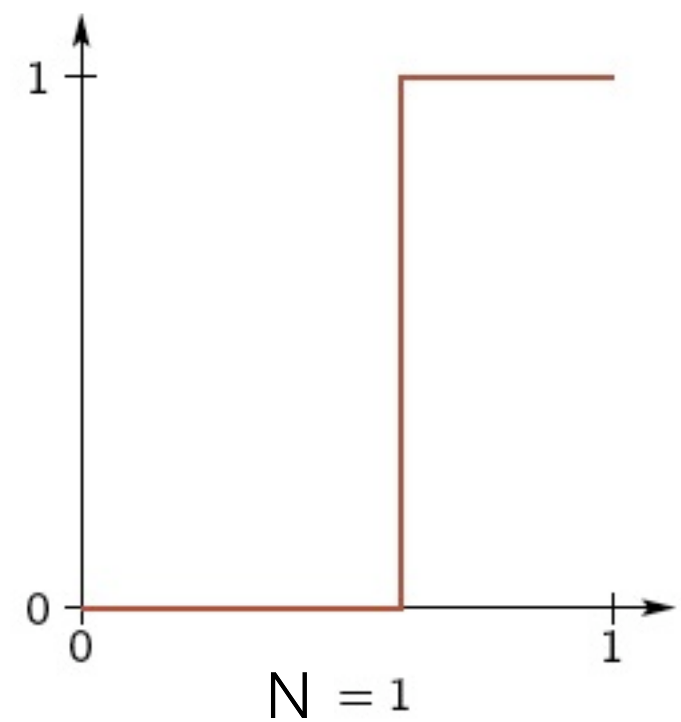
$$\begin{aligned} I(P^-) + I(P^+) &= I(U_1 U_2; Y_1 Y_2) \\ &= I(X_1; Y_1) + I(X_2; Y_2) \\ &= 2I(P), \end{aligned}$$

$$I(P^-) \leq I(P) \leq I(P^+)$$

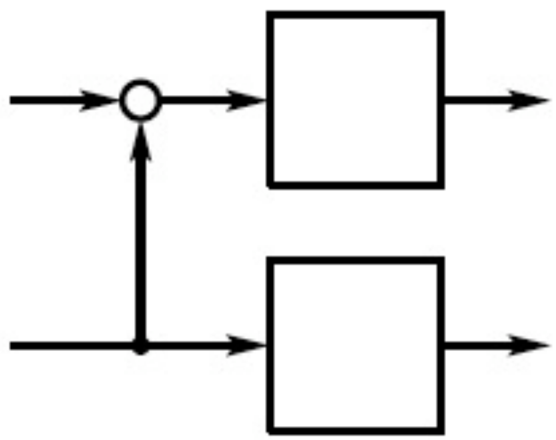
More on Polarization

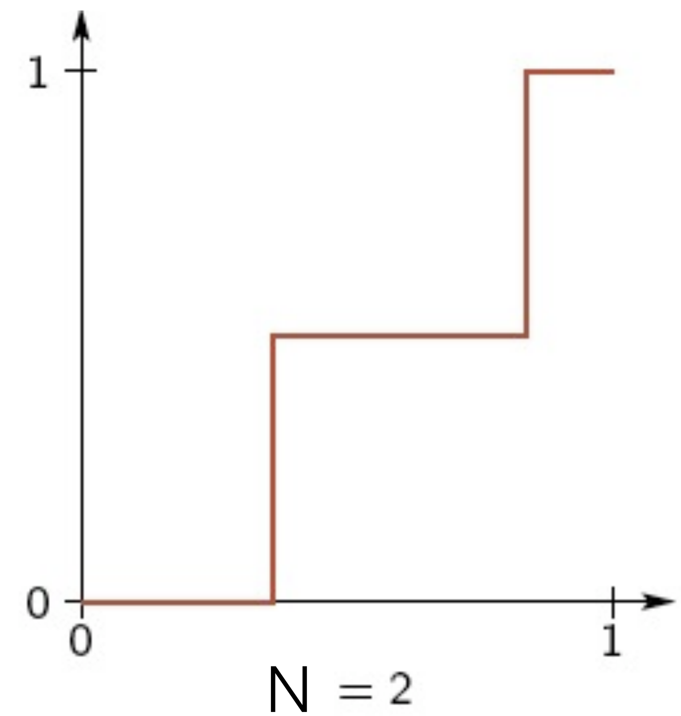
More on Polarization



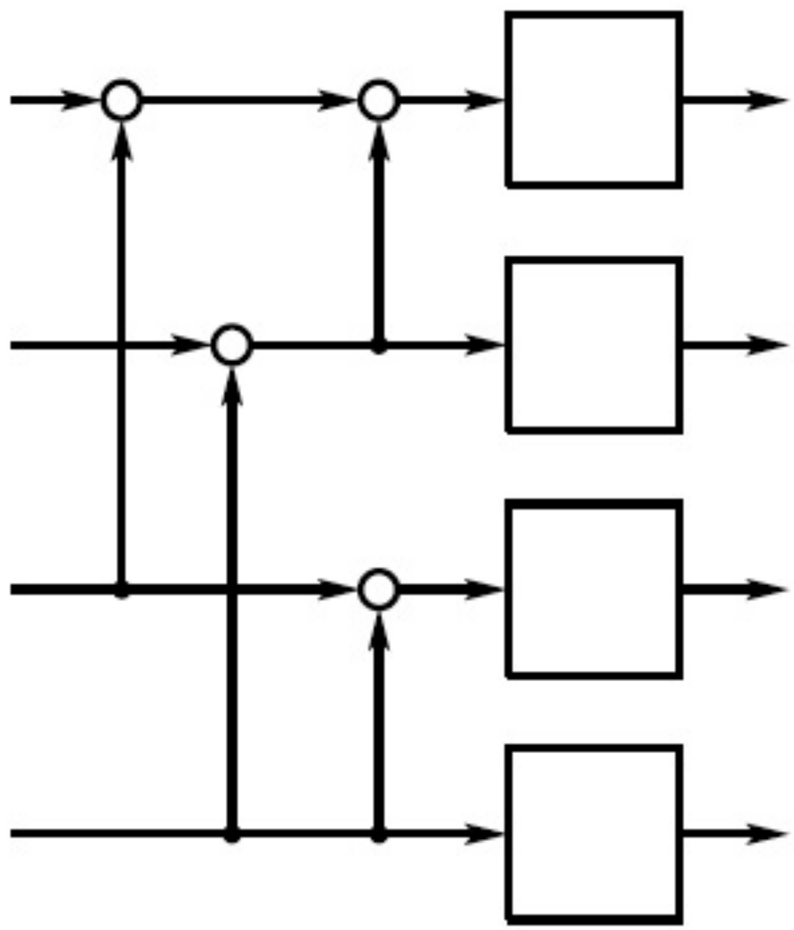


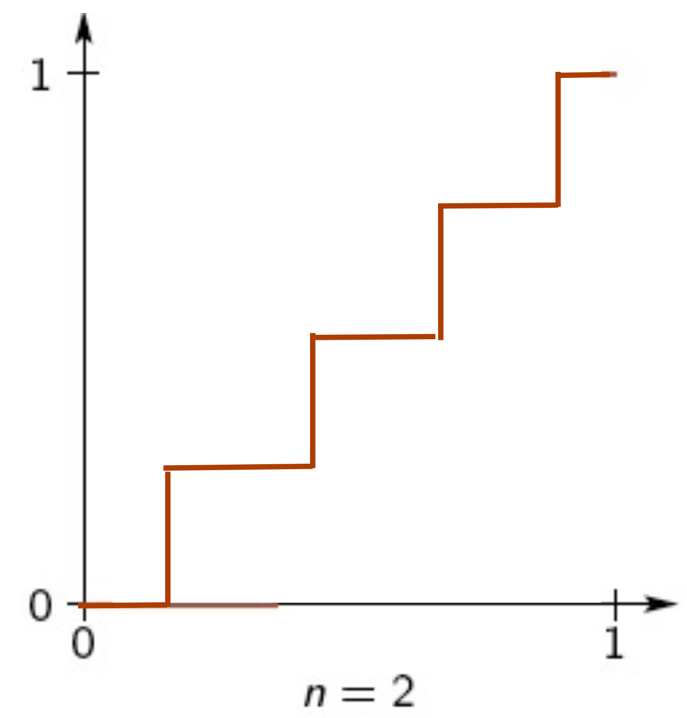
More on Polarization



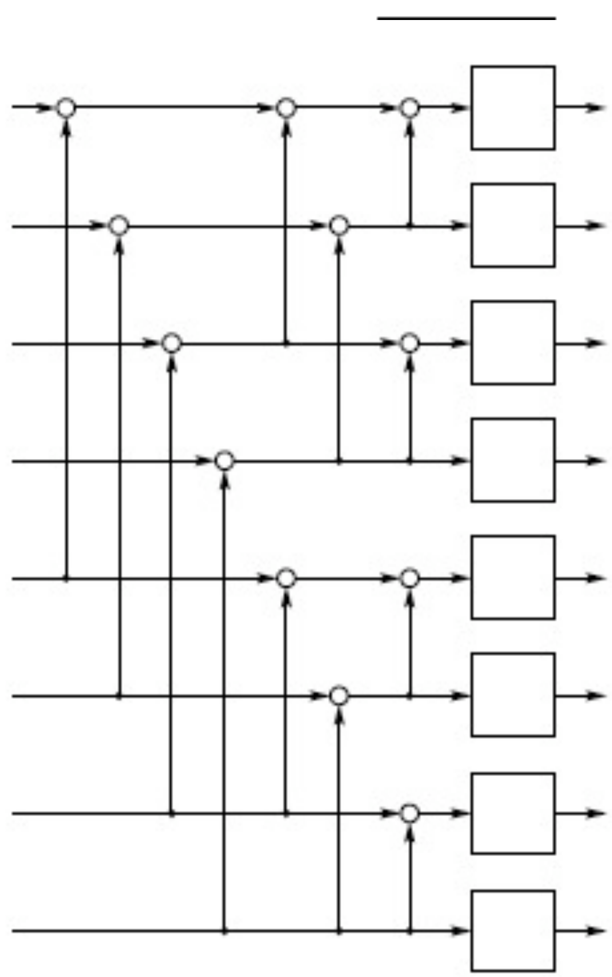


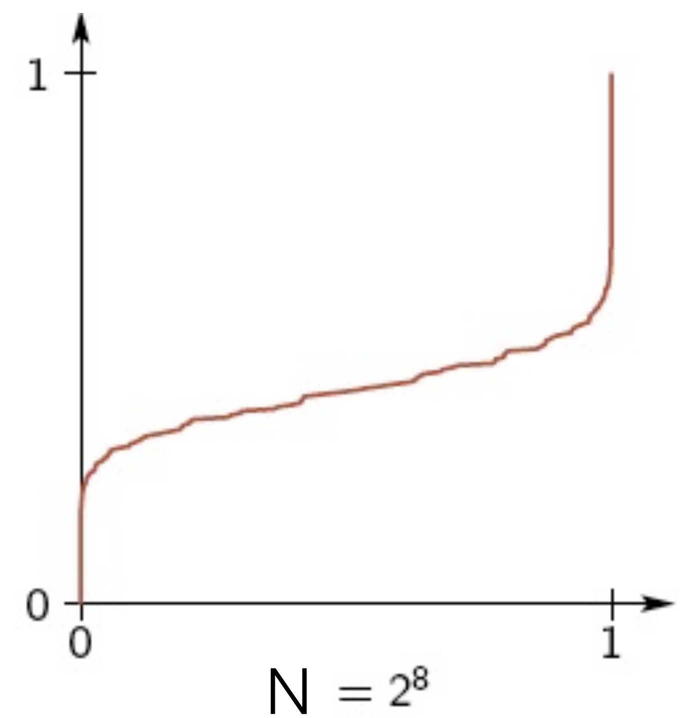
More on Polarization





More on Polarization





Equivalent “Random” Channel

Equivalent “Random” Channel

Set B1, B2, ... to be i.i.d. {+, -} valued,
uniformly distributed random variables

Equivalent “Random” Channel

Set B1, B2, ... to be i.i.d. {+, -} valued, uniformly distributed random variables

Define $I_n = I(W^{B_1, B_2, ..., B_n})$

Equivalent "Random" Channel

Set B1, B2, ... to be i.i.d. {+, -} valued, uniformly distributed random variables

Define $I_n = I(W^{B_1, B_2, \ldots, B_n})$

Study the distribution of I_n

Properties of I_n

Properties of I_n

$I_0 = I(W)$ is a constant

Properties of I_n

$I_0 = I(W)$ is a constant

$I_n \in [0, 1]$; so I_n is bounded

Properties of I_n

$I_0 = I(W)$ is a constant

$I_n \in [0, 1]$; so I_n is bounded

Conditional on $B_1, B_2, ..., B_n$, and with $P = W^{B_1, B_2, ..., B_n}$, I_{n+1} can only take on the two values $I(P^+)$ and $I(P^-)$

Properties of I_n

$I_0 = I(W)$ is a constant

$I_n \in [0, 1]$; so I_n is bounded

Conditional on $B_1, B_2, \ldots, B_n$, and with $P = W^{B_1, B_2, \ldots, B_n}$, I_{n+1} can only take on the two values $I(P^+)$ and $I(P^-)$

Further, $E[I_{n+1} \mid B_1, B_2, \ldots, B_n] = (I(P^+) + I(P^-))/2 = I(P)$, so $\{I_n\}$ is a (bounded) martingale

Properties of I_n

Properties of I_n

a bounded martingale converges almost surely

Properties of I_n

a bounded martingale converges almost surely

$I_\infty = \lim_{n \to \infty} I_n$ exists almost surely; $E[I_\infty]=I_0=I(W)$

Properties of I_n

a bounded martingale converges almost surely

$I_\infty = \lim_{n\to\infty} I_n$ exists almost surely; $E[I_\infty]=I_0=I(W)$

$\Pr\{|I_{n+1}-I_n|\leq\varepsilon\}\to 1$; but $|I_{n+1}-I_n|=(I(P^+)- I(P^-))/2$

Properties of I_n

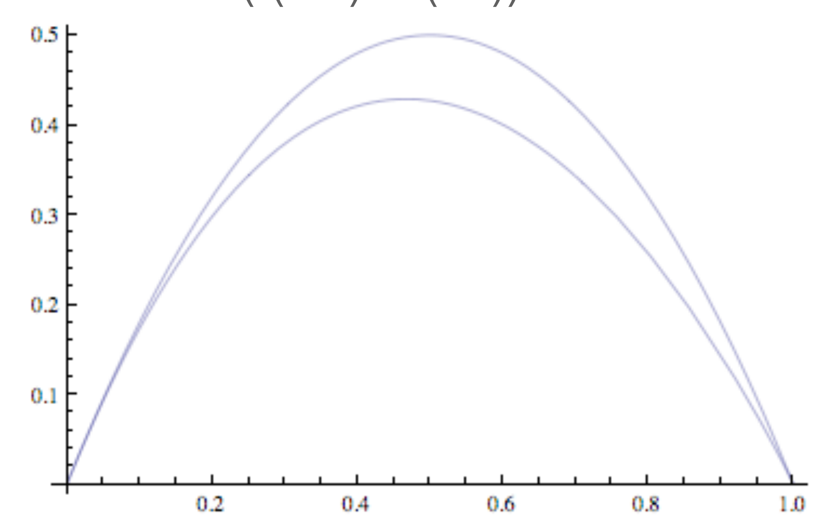
a bounded martingale converges almost surely

$I_\infty = \lim_{n\to\infty} I_n$ exists almost surely; $E[I_\infty]=I_0=I(W)$

$\Pr\{|I_{n+1}-I_n|\le\varepsilon\}\to 1$; but $|I_{n+1}-I_n|=(I(P^+)- I(P^-))/2$

from extremes of information combining we know that $(I(P^+)-I(P^-))/2 \le\varepsilon$ implies that $I(P)\notin(\delta, 1-\delta)$

(I(P+)- I(P-))/2



I(P)

Properties of I_n

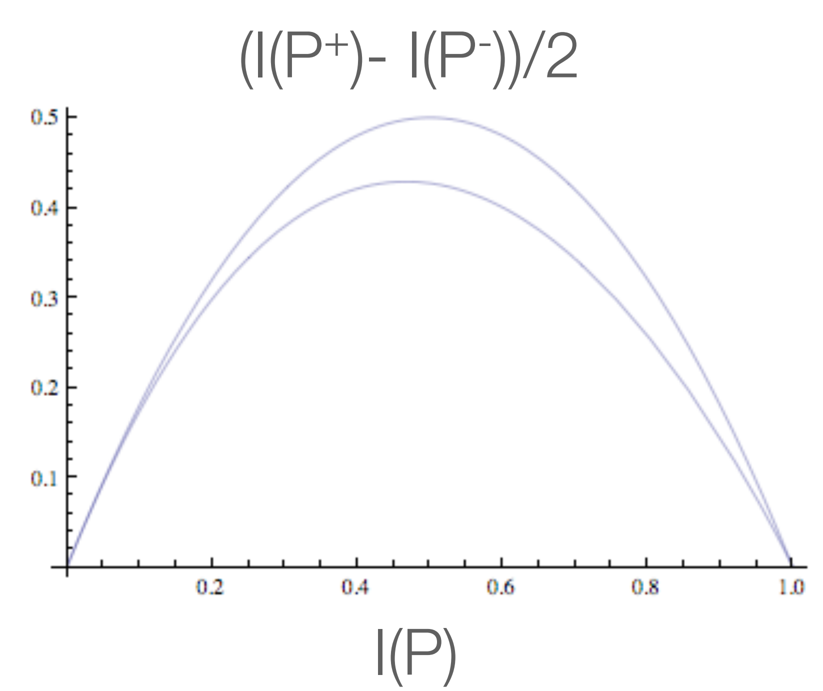
a bounded martingale converges almost surely

$I_\infty = \lim_{n\to\infty} I_n$ exists almost surely; $E[I_\infty]=I_0=I(W)$

$\Pr\{|I_{n+1}-I_n|\leq\varepsilon\}\to 1$; but $|I_{n+1}-I_n|=(I(P^+)-I(P^-))/2$

from extremes of information combining we know that $(I(P^+)-I(P^-))/2 \leq \varepsilon$ implies that $I(P)\notin(\delta, 1-\delta)$

we conclude that I_∞ takes values only in $\{0, 1\}$



Summary of Known Results

Summary of Known Results

achieve capacity on memoryless channels   Arikan 2007

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{Ec(G)}}$	Korada, Sasoglu, and U. 2009

Polar Codes Based on Larger Matrices

In [Arıkan08] generator matrix is constructed from the rows of $G_2^{\otimes m}$, where

$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Polar Codes Based on Larger Matrices

In [Arıkan08] generator matrix is constructed from the rows of $G_2^{\otimes m}$, where

$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Consider codes constructed from $G^{\otimes m}$, where G is an $\ell \times \ell$ matrix. Blocklength $N = \ell^m$, $\bar{x} = \bar{u}G^{\otimes m}$

Characterization of Exponent

Theorem (Korada,Şaşoğlu,Urbanke)

For any $R < I(W)$, and $\ell \times \ell$ matrix G,

$$P_B(N) \approx 2^{-(N)^{\mathtt{E}_c(G)}},$$

where $\mathtt{E}_c(G) = \frac{1}{\ell}\sum_{i=1}^{\ell} \log_\ell d_i$.

$$G = \begin{bmatrix} g_1 \\ \vdots \\ g_i \\ g_{i+1} \\ \vdots \\ g_\ell \end{bmatrix} \qquad d_i = \mathtt{dmin}(g_i, \langle g_{i+1}, \ldots, g_\ell \rangle)$$

Exponent: Example

Exponent: Example

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Exponent: Example

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$d_1 = 1,$$

Exponent: Example

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$d_1 = 1, d_2 = 2,$$

Exponent: Example

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$d_1 = 1,\ d_2 = 2,\ d_3 = 2$$

Exponent: Example

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$d_1 = 1,\ d_2 = 2,\ d_3 = 2$$

$$\mathtt{E}_c(G) = \frac{1}{3}(\log_3 1 + \log_3 2 + \log_3 2) = 0.42062$$

Results

Results

$$\mathtt{E}_\ell = \max_{G \in \{0,1\}^{\ell \times \ell}} \mathtt{E}_c(G)$$

Results

$$\mathtt{E}_\ell = \max_{G \in \{0,1\}^{\ell \times \ell}} \mathtt{E}_c(G)$$

For $\ell < 15$, $\mathtt{E}_\ell \leq \frac{1}{2}$

Results

$\mathbf{E}_\ell = \max_{G \in \{0,1\}^{\ell \times \ell}} \mathbf{E}_c(G)$

For $\ell < 15$, $\mathbf{E}_\ell \leq \frac{1}{2}$

$\mathbf{E}_{16} = 0.51828$

Results

$\mathtt{E}_\ell = \max_{G \in \{0,1\}^{\ell \times \ell}} \mathtt{E}_c(G)$

For $\ell < 15$, $\mathtt{E}_\ell \leq \frac{1}{2}$

$\mathtt{E}_{16} = 0.51828$

$\lim_{\ell \to \infty} \mathtt{E}_\ell = 1$

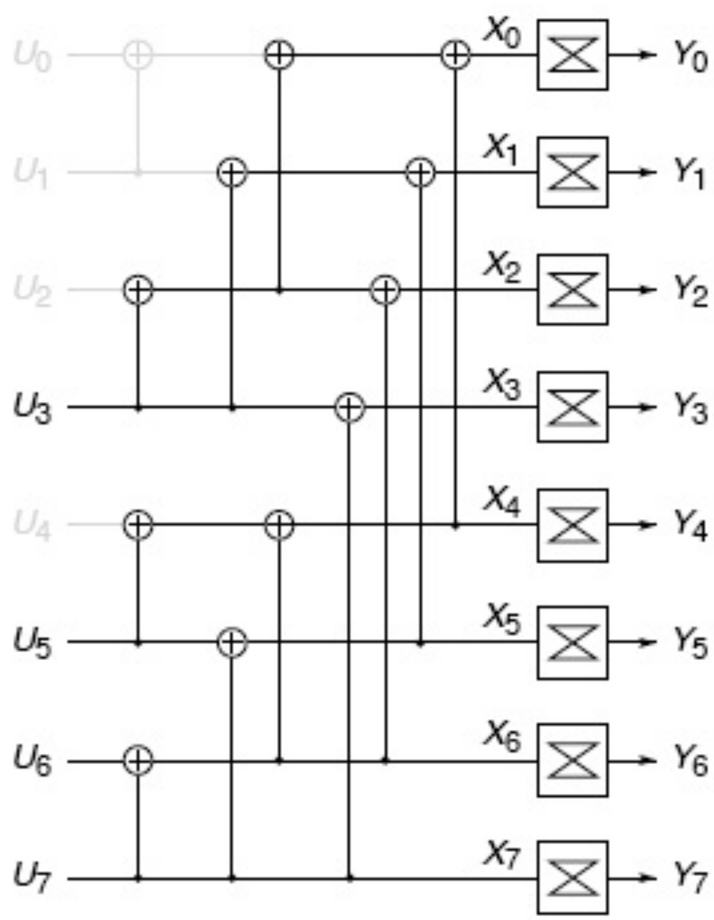
Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{Ec(G)}}$	Korada, Sasoglu, and U. 2009

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{Ec(G)}}$	Korada, Sasoglu, and U. 2009
optimal for lossy source coding, Wyner-Ziv, Gelfand-Pinsker, ...	Korada and U. 2009

Source Coding



Y-binary symmetric source

$\mathtt{d}(0,1) = 1$, $\mathtt{d}(0,0) = \mathtt{d}(1,1) = 0$

$\bar{Y}$- source word, $\bar{X}$-reconstruction wor

$\frac{1}{N}\mathbb{E}[\mathtt{d}(\bar{X}, \bar{Y})] \leq D$

$R > 1 - h_2(D)$

Source Coding

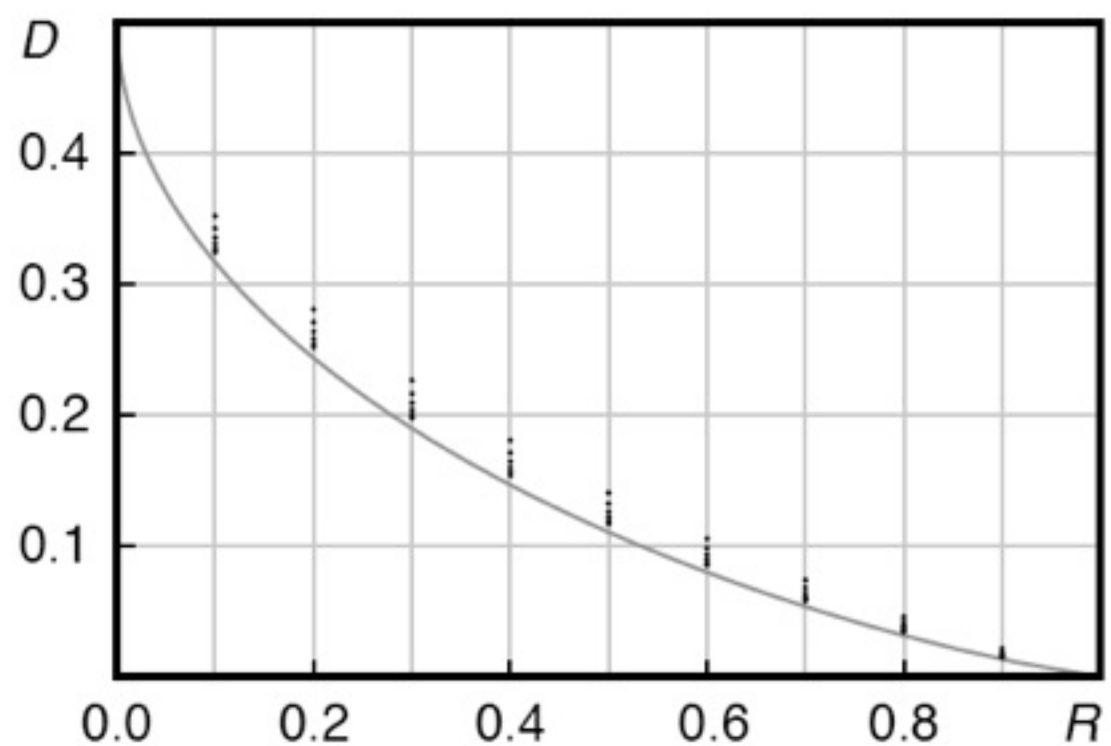
- trellis based quantization [Viterbi and Omura], constraint length to infinity
- LDGM based quantization [Martinian and Yedidia], [Wainwright and Maneva], [Ciliberti and Mézard], [Filler and Fridrich], works well in practice
- low-density compound constructions [Wainwright, Martinian], uses MAP

Source Coding

trellis based quantization [Viterbi and Omura], constraint length to infinity

LDGM based quantization [Martinian and Yedidia], [Wainwright and Maneva], [Ciliberti and Mézard], [Filler and Fridrich], works well in practice

low-density compound constructions [Wainwright, Martinian], uses MAP



Length = $2^9, 2^{11}, \ldots, 2^{19}$

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{E_c(G)}}$	Korada, Sasoglu, and U. 2009
optimal for lossy source coding, Wyner-Ziv, Gelfand-Pinsker, ...	Korada and U. 2009

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{E_c(G)}}$	Korada, Sasoglu, and U. 2009
optimal for lossy source coding, Wyner-Ziv, Gelfand-Pinsker, ...	Korada and U. 2009
efficient construction	Mori and Tanaka 2009

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{E_c(G)}}$	Korada, Sasoglu, and U. 2009
optimal for lossy source coding, Wyner-Ziv, Gelfand-Pinsker, ...	Korada and U. 2009
efficient construction	Mori and Tanaka 2009
suboptimal for compound coding	Hassani, Korada and U. 2009

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{E_c(G)}}$	Korada, Sasoglu, and U. 2009
optimal for lossy source coding, Wyner-Ziv, Gelfand-Pinsker, ...	Korada and U. 2009
efficient construction	Mori and Tanaka 2009
suboptimal for compound coding	Hassani, Korada and U. 2009
non-binary version and asym. channels	Arikan, Sasoglu, and Telatar 2009

Summary of Known Results

achieve capacity on memoryless channels	Arikan 2007
$P_B(N) \approx 2^{-N^{\frac{1}{2}}}$	Arikan and Telatar 2008
$P_B(N) \approx 2^{-(N)^{E_c(G)}}$	Korada, Sasoglu, and U. 2009
optimal for lossy source coding, Wyner-Ziv, Gelfand-Pinsker, ...	Korada and U. 2009
efficient construction	Mori and Tanaka 2009
suboptimal for compound coding	Hassani, Korada and U. 2009
non-binary version and asym. channels	Arikan, Sasoglu, and Telatar 2009
scaling	

Summary

+ completely new paradigm of coding

+ provably achieves capacity

+ low complexity

+ many applications

- currently only competitive for VERY large N

Wyner-Ziv and Gelfand-Pinsker

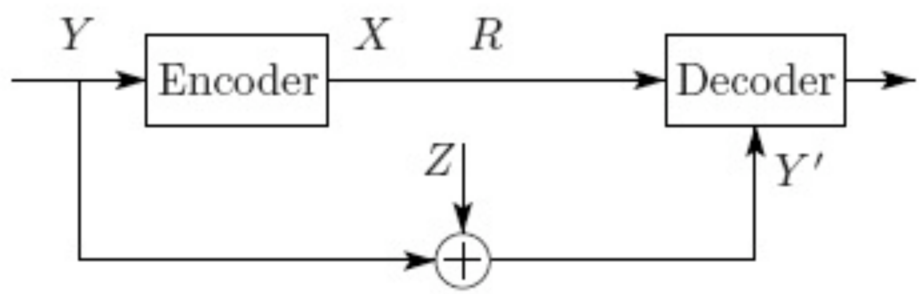


Figure 4.1: The Wyner-Ziv problem. The task of the decoder is to reconstruct the source Y to within a distortion D given (Y', X).

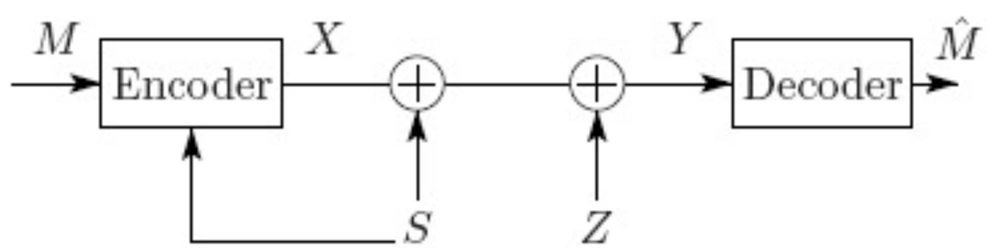


Figure 4.4: The Gelfand-Pinsker problem. The state S is known to the encoder a-causally but not known to the decoder. The transmission at the encoder is constrained to have only a fraction D of ones on average, i.e., $\mathbb{E}[X] \leq D$.